

Farmers and conservationists

The long-standing conflict between agriculturists and conservationists, recently sharpened in Britain, will engulf Europe in the next few years. The British House of Lords has pointed comment.

THE great European row between farmers and conservationists, which has been simmering away for years, is coming quickly to the boil. Part of the explanation is that Europe knows that it is foolish to subsidize farmers to produce food that cannot be absorbed in Europe or sold elsewhere at any knock-down prices. This year's reluctant agreement among the member states of the European Communities that surplus production of dairy products must be cut back by means of a quota system rather than by a price reduction, economic nonsense though it is, has nevertheless so outraged the farming lobbies that the taxpayers who ultimately provide the subsidies have begun to rebel: if we have to pay towards the countryside, they argue, why cannot we have some say in how it looks? Further fuel has been added to the fire by the discovery in the United Kingdom that the Wildlife and Countryside Act of 1981, thought to provide the Nature Conservancy Council with the legal apparatus for protecting natural sites of particular importance, it at best a leaking vessel. The row between farmers and conservationists has now been brought into the open by the publication, earlier this week, of a report of the House of Lords Select Committee on the European Communities on Agriculture and the Environment (HL 247).

The House of Lords Committee, as always, has done a thorough job in exposing the contradictions embodied in the agricultural policies of the European Commission and of the member governments of the Communities. The occasion is the publication last year of the Commission's latest directive on the provision of aid to European farmers. Already alarmed that the accumulation of farm surpluses threatened its own survival, the Commission proposed a general realignment of agricultural subsidies intended to reduce the supply of commodities in surplus. At the same time, but in flat contradiction, it proposed that governments should be free to pay capital grants for the improvement of farms even when these are too small to support their owners' families. Not surprisingly, the House of Lords concludes that the new regime would be hardly less encouraging to food production than the present. But it also finds that in the United Kingdom, both the agriculture and the environment departments take such a narrow view of the Treaty of Rome that they have been less clever at blending the objectives of production and conservation than the European Commission would allow (and even pay for).

In Britain, as in most of the rest of Europe, the underlying difficulty is that agriculture departments are the captives of their farmers, but unthinkingly. The explanation is that the underlying objective is to make farmers economically viable. What the House of Lords points out is that the European Communities provide for the payment of grants to farmers in what are called "less favoured areas", uplands or coastal marshes, for example, which are not strictly geared to increasing agricultural production. So it would be possible for the British Government to assist the economic prosperity of rural areas not only by paying grants for the drainage of unstable hillsides, or for building roads cutting vertically up hillsides, but also for establishing small woodlands and by encouraging rural tourism. The ministry witnesses before the House of Lords committee seem dimly to have understood that there was some objection to the ministry's traditional practice, but not to have understood that a little cleverness might allow it to do more for its pensioners and at the same

time to safeguard vulnerable landscape.

The immediate difficulty is that the negotiation of the final form of the new directive will be the responsibility not of the House of Lords but of the ministries of agriculture. The conflict between agricultural production and conservation is, however, too important to be left to them. The ideal solution, probably politically unattainable, is that the Communities' support prices for food should be adjusted downwards until there is a balance, taking one year with another, between supply and demand, and that farming communities put in economic jeopardy in the process should be helped from the Communities' social fund.

Short of that solution, there is much that governments could do to help themselves. In Britain it is folly to make annual payments to farmers for each head of stock in their possession (as is the present practice in upland areas) even when the result is a damaging degree of overgrazing. Similarly, it is foolish that the system for paying grants for "farm improvements" should most often (except in national parks and other special areas) require no previous scrutiny on conservation grounds. While such practices continue, there is no hope that the relationship between farmers and conservationists can be anything but hostile.

The peculiar nonsense of the British Countryside Act is a further invitation to inflamed relationships. For more than thirty years, the Nature Conservancy Council has been trudging around the United Kingdom, designating certain sites as "sites of special scientific interest". Sometimes they may be found to carry a distinctive flora, or to provide a special habitat for wildlife. Many sites have been so designated simply because of their natural beauty. Some have distinctive geological interest. For much of its history, the council has done its best to preserve these sites intact by means of voluntary management agreements with the landowners. The Countryside Act in 1981 sought to bring these arrangements within a more robust framework by giving the council powers to prevent the agricultural or other development of these sites — and the duty to compensate landowners if agreement could not be reached.

What has happened now is that landowners have begun sending in their bills for refraining from the development of designated sites. Nobody can tell whether a proposed development is seriously intended or merely a threat. A strange irony of the arrangements laid down by the government is that a landowner's compensation will be calculated after taking account of the agricultural subsidies he would receive if development were allowed. The signs are that the annual demand for compensation will quickly outstrip what the council is allowed by the government to afford. Then the fat will be thoroughly in the fire.

Before that happens, the British Government should move swiftly to redress the balance between farmers (who are often needlessly maligned) and the conservationists (often their own worst enemies). In Britain, where agriculture has been almost entirely exempted from the strict planning laws that prevent many people from cutting down a tree or erecting a garden shed, the disciplines of seemly neighbourliness need to be carried into the countryside. With an eye on the future, there should also be a deliberate retreat from agricultural production on the most marginal of marginal land, at least at the taxpayers' expense. And somebody should make sure that the ministries of agriculture mean what they say with their lip-service to conservation. □



US astronomy

Kitt Peak and other observatories shaken up

Washington

KITT Peak National Observatory is about to complete its uneasy transition to a new management regime intended to accelerate work on the next-generation large optical telescope, reemphasize an in-house research programme and save money into the bargain. The changes have also resulted in the easing out of the director of Kitt Peak, Geoffrey Burbidge, who just a year and a half ago had been given a five-year renewal of his contract. The decision to replace Burbidge has raised more than a few eyebrows among users of the observatory, who had generally been pleased with his efforts to serve the needs of visiting researchers.

The chief structural change in the management has been unification of the three observatories run by the Association of Universities for Research in Astronomy, Inc. (AURA), under a single "superdirector". The three — Kitt Peak, Cerro Tololo Inter-American Observatory and the National Solar Observatories (Sacramento Peak and McMath Observatories) — are to have a single unified scientific staff as well, a change that will rejuvenate the observatories' languishing in-house research programme and facilitate the transfer of staff among the observatories.

John Jefferies, formerly director of the Institute for Astronomy at the University of Hawaii, is the new superdirector, with the formal title of director of the National Optical Astronomy Observatories (NOAO). Jefferies, whose strong management is credited with building up the Hawaii institute to a major research facility, has moved quickly since his appointment to reshape NOAO along his lines. Last January, AURA announced that since the roles of the three observatory directors would change "substantially" under Jefferies' planned reorganization, they would be considered "new positions", so that existing contracts were void. Both Burbidge and Jack Zirker, director of the solar observatories, had had their appointments renewed for five-year terms in 1982; both found their jobs advertised earlier this year. Neither was selected. In Burbidge's place, Sydney Wolff, Jefferies' former assistant at Hawaii, was appointed. Zirker was replaced by Robert Howard, the Carnegie Institution's former principal solar investigator. Patrick Osmer, the director of Cerro Tololo, was asked to remain in post; the reason given was that he had two years remaining of his term.

Although Burbidge was well liked by visiting astronomers at Kitt Peak, he was, in the word of one astronomer familiar with the situation, "hated" by the on-site

staff. Burbidge did an about-face from the policy of the previous director by insisting that in the face of limited funding, the staff's responsibility was to serve the research community first and its own research interests second. AURA is supported by the National Science Foundation (NSF) to the tune of approximately \$25 million a year.

There was also said to be dissatisfaction with the pace of work on the National New Technology Telescope (NNTT), which had been designated a top priority in astronomy by the National Academy's so-called "Field Report". This telescope would be the next major optical facility and would employ either multiple-mirror or segmented-mirror technology. Jefferies created a new division, on a par with the three observatories, to take over planning for NNTT.

The desire to step up planning for NNTT may indeed be the chief reason behind the entire shake-up at Kitt Peak. Jefferies has built a reputation on his ability to get big projects off the ground. And indeed, the management rationale for the creation of the superdirectorship laid out in Jefferies' own plan, dated December 1983, wears a bit thin. Apart from making NOAO into "an effective and efficient organization", the superdirector is to engender "a sense of unity and purpose ... through his personal presence, example and influence".

NSF is pleased with the changes, according to Laura P. Bautz, director of NSF's astronomical sciences division; NSF hopes the pooling of the administrative functions of the three observatories under the superdirector may lead to greater efficiency.

One other thing that the superdirectorship will do is take the heat off of AURA's board of directors with respect to dividing up the money from NSF among the various observatories. Dr John Teame, president of AURA, had had that responsibility; Teame said it had been felt that it would be much better to have those decisions made by somebody in a position of clear scientific leadership. **Stephen Budiansky**

Mt Wilson telescope for axe

Washington

AFTER months of rumours, the Carnegie Institution of Washington has announced that it plans to cease operating the 100-inch reflector telescope on Mt Wilson on 1 July next year. The decision is part of an overall plan by the institution to shift observational work to its Las Campanas observatory in Chile; it also reflects the diminishing importance of Mt Wilson as the lights of nearby Los Angeles get ever brighter. The institution operates 40-inch and 100-inch reflectors at Las Campanas.

Carnegie also plans a "phased reduction" in support for two solar telescopes on Mt Wilson. The 60-inch reflector, which is being used in a systematic study of Sun-like stars that has yielded valuable data on solar radiation periods, will continue in operation. That project is supported largely by National Science Foundation (NSF) grants, which are expected to continue for several years.

Dr George Preston, director of the observatories, noted the need to upgrade computer facilities and detectors at Las Campanas. Carnegie, which relies on a hefty endowment income supplemented by government and foundation grants, spent \$1.5 million to operate Mt Wilson last year and \$900,000 to operate Las Campanas.

Despite some optimistic statements from Carnegie officials about finding an organization to take over the Mt Wilson telescopes, few astronomers seem to think that likely. The 100-inch reflector has been used mainly by on-site staff in recent years; and because of the light pollution the work is limited to observing single relatively-

bright stars. "Only a very restricted class of programs can be taken seriously at that site" said Rudy Schild of the Harvard-Smithsonian Astrophysical Observatory. The more exciting observational work on extra-galactic objects requires the dark-sky conditions now available only in the Southern Hemisphere.

Lee Hartmann, another astronomer at Harvard-Smithsonian who has used Mt Wilson, says that the only way to "sell" the 100-inch reflector would be another solar-stellar programme like that in progress at the 60-inch Mt Wilson telescope. That would be "politically difficult" to justify, he said. It is generally considered to be out of the question that the most obvious supporters of astronomical research — NSF and the National Aeronautical and Space Administration — would be interested in taking over the 100-inch telescope. Interest in the solar telescopes is said to run higher. Carnegie officials refuse to discuss the status of negotiations that they say are now taking place with various interested parties. **Stephen Budiansky**

Correction

IN the 7 June issue of *Nature* (p.500), reference is made in the penultimate paragraph to a "Shell and Gist Brocades . . . joint, DfL300 million, biotechnology laboratory", which is not correct. Shell is involved in a joint research venture with Gist Brocades. However, all the research is carried out at the (existing) Shell Sittingbourne Research Centre in the United Kingdom and at Gist Brocades' laboratory in Delft, the Netherlands.

US research budget

Spending plan remains intact

Washington

CONGRESS left town last week for the first of several long recesses in this election year with a respectable share of work on the budget for the financial year beginning in October completed.

The administration has suffered few setbacks in its overall plan to increase basic science and military research, emphasizing the "hard" sciences and a few special pet projects — such as the Center for Advanced Materials at Lawrence Berkeley Laboratory. Congress has promoted a few pet projects of its own while keeping to a minimum its more public-spirited tinkering with the budget.

NSF: The National Science Foundation (NSF) will receive \$1,302 million for research and related activities, only \$6 million short of the administration request. Congress doubled the \$20 million that the Reagan Administration wanted for supercomputers, opening the way for the purchase of new machines or the establishment of national supercomputer centres — rather than simply facilitating the use of existing machines, as the administration had planned. The House-Senate compromise on the final version of the NSF bill made a special point of calling on NSF not to establish any such centre without competitive review of proposals — apparently a reaction to less high-minded handling of supercomputer centres in the Department of Energy budget (see below).

Congress also directed NSF to cut \$36 million from the programmes that the Reagan Administration has tended to favour in the past few years — physical sciences and engineering — at the expense of biological, behavioural and social sciences.

At the urging of Congress, outgoing NSF director Ed Knapp provided written assurance that 13,000 grants will be awarded next year.

Congress increased NSF's science education budget, a perennial target of the Reagan Administration, and in particular moved to maintain the present number of 600 NSF graduate fellowships. The annual stipend was increased to \$11,000 plus a tuition allowance of \$6,000.

Energy: The administration was only slightly less successful in getting its way in the energy budget. Congress for the most part went along with the Reagan plan to place less emphasis on applied research, particularly that concerned with solar and renewable energy, while building up basic research. In one area, magnetic fusion, Congress went one better than the administration, cutting that programme to \$420 million — an 11 per cent drop from present levels.

One setback to the administration was the refusal of Congress to appropriate construction funds for the planned

electron accelerator to be built at Newport News, Virginia, by the Southwestern Universities Research Association (see *Nature* 21 June, p.658). In place of construction funds, Congress allocated \$3.5 million for research and development and preliminary design work. Before coming back to Congress for construction funds, the Department of Energy will have to submit a five-year plan on nuclear physics that examines the possibility of dropping the project altogether.

The two projects that Congress inserted last year into the energy budget without the benefit of peer review — and which gave rise to cries of pork-barrel politics (see *Nature* 305, 659; 1983) — are back again; Catholic University will get \$9 million to continue work on its Vitreous State Laboratory, and Columbia University \$3 million for its "National Center for Chemical Research", also known as a chemistry building. Not to be left out, Representative Don Fuqua (Democrat, Florida), chairman of the House Science and Technology Committee, saw to it that \$7 million was allocated for a new supercomputer centre at Florida State University, in his home state.

The Center for Advanced Materials, presidential science adviser George Keyworth's showcase for government-industry cooperation through the national laboratories, will get \$11 million in construction funds as requested. Work on the Stanford Linear Collider will continue as planned.

Space: The National Aeronautics and Space Administration (NASA) will be given \$10 million more than the administration had intended for research analysis — \$7 million of which will go for planetary research and analysis — a modest response to the frequently-heard criticism that

NASA has promoted new projects at the expense of budgets needed to analyse data piling up from existing projects.

Congress agreed to provide \$150 million for research and development on the proposed space station, but in turn required that NASA should consider, as one alternative, that permanent manning of the station be phased in over a five-year period following the station's deployment. The Venus Radar Mapper, the Galileo mission to Jupiter, the gamma-ray observatory and the space telescope will proceed as planned.

Other: The Environmental Protection Agency will receive a big increase in its research and development budget — \$193 million, a 35 per cent increase over the 1984 level. The administration had requested \$163 million for fiscal year 1985.

Neither the House nor the Senate has acted on the National Institutes of Health (NIH) appropriation. The Senate Appropriations Committee is expected to act quickly when Congress returns at the end of the month with a recommendation to increase the administration's request by \$587 million. The administration wanted \$4,342 million for NIH's research activities.

The agriculture budget, which was to have included a major new initiative on biotechnology as a vehicle for expanding the perpetually minuscule competitive grants programme, was given its usual treatment in the House by Appropriations Committee chairman Jamie Whitten (Democrat, Mississippi) (see *Nature* 14 June, p.574). The result was that while the overall total of research funds has changed very little from the administration's proposed 1 per cent increase over 1984, the competitive grants programme was once again slashed. The Senate, which has yet to act, is expected to look much more favourably on the administration's plan to bring the competitive grants programme up to \$50 million, from its current \$17 million.

Stephen Budiansky

Soviet emigré case not resolved

PARTICIPANTS returning from the Federation of European Biochemical Societies (FEBS) Moscow meeting have reported further details about the case of Dr David Goldfarb, the Jewish biochemist (see *Nature*, 308, 766; 309, 104; 1984). Informants who were able to meet Dr Goldfarb say that, although he was promised an exit visa in February, he never received it, but was still waiting for the relevant documents at the time of the police search of his apartment in April. On the bacterial strains confiscated by the police, he said that he had already decided not to take them with him.

Unlike most "refusniks", Dr Goldfarb was still employed by a research institute until notified that he would receive a visa.

When Goldfarb resigned from the institute this spring, he took with him

examples of the strains he had developed in the 1960s for urine tests to determine metabolic disorders in children. He then decided, however, not to take the strains abroad with him, since this might create difficulties at the customs and because the strains had, in any case, been obtained from the United States and the United Kingdom. He says, however, that during his interrogation, a KGB official stated that he was under investigation both for trying to take the strains out of the country and for disseminating hostile propaganda.

Dr Goldfarb's present status is unclear. Many of the FEBS participants say they raised the subject with Academician Yuri Ovchinnikov and received encouraging answers. However, Ovchinnikov seems to have given no definite commitment that Goldfarb would receive a visa. Vera Rich

UK nuclear tests

Another Australian inquiry

Canberra

GROWING public disquiet that either the Australian or British Government is concealing politically sensitive atmospheric nuclear fallout data has prompted the Australian Minister for Resources and Energy, Senator Peter Walsh, to recommend that a Royal Commission be set up to inquire into the conduct of the British nuclear tests in Australia between 1952 and 1963. This move follows a procession of government-sponsored investigations into the tests, with each successive report revealing discrepancies in the last, and punctuated in recent months by sensational newspaper revelations, the most startling being the claim from an anonymous source that mentally-retarded adults were used as guinea pigs in tests at Maralinga in 1956.

On 31 May, an expert committee headed by Professor Charles Kerr, which had only sixteen days to review the available fallout data, reported to Senator Walsh that it was extremely critical of what it saw as obfuscation, oversimplified assumptions and a tendency to present best-case interpretations of data in the January 1983 report of the Australian Ionizing Radiation Advisory Council (AIRAC 9). The Kerr committee had used the AIRAC 9 document in its review of data collection methodology, but could not dismiss the possibility that the "black mist" described by the Pitantjatjara aboriginal people may have represented a deviant fallout pattern from a test at Emu in 1953. It also cast doubt on the statistical assumptions underpinning a November 1983 Department of Health study which rejected associations between illness and irradiation of personnel at the test sites.

On 7 June, Senator Walsh tabled another document, a chronology of the British tests prepared by Dr J.L. Symonds and gathered mainly from official Australian sources. Among many other facts never officially admitted before, it made clear that the second, very dirty, trial of the 1956 Monte Bello series, Mosaic G2, had involved lighter atomic elements and could be described as an H-bomb trigger. (The UK H-bomb test took place at Christmas Island the following year.) The chronology indicates that the Australian Government was well aware that the yield would be 60 kilotons, yet in 1983 AIRAC 9 reported the government's 1956 claim that the yield was "in the kiloton range". The hundreds of minor trials conducted at Maralinga went virtually unreported in AIRAC 9. The Symonds chronology, on the other hand, highlighted the deeply-felt Australian ignorance of the minor trials known as Vixen A and B, which were meant to simulate the accidental detonation of plutonium weapons due to damage or misfiring, and could result in a

nuclear blast. Begun during the 1958-61 moratorium, the Vixen tests were discontinued in 1963 since, as Symonds reports, they might have been classified as violations of the test ban.

Soon, however, gaps were to be found in the Symonds report. Later in June, newspaper stories claimed that the specially-prepared warship HMS *Diana* had sailed deliberately into a fallout cloud after a 1956 Monte Bello explosion, and that mentally-retarded people had been placed in bunkers one mile from one of the 1956 Buffalo series of blasts at Maralinga, a story later amplified by Mr Terry Toon, the national secretary of the Maralinga and Monte Bello Atomic Ex-Servicemen's Association, who

so far has not named the others able to corroborate his claim to have heard a scream from one of the forward bunkers.

On 5 July, Senator Walsh announced the Royal Commission, which will focus on the measures taken to protect personnel at the sites and aboriginals in the surrounding area. The British Government has agreed to cooperate fully and to guarantee immunity from prosecution for ex-servicemen and civilians who give evidence in Australia. In dealing with five compensation claims already launched against it, the Australian Government has waived the statute of limitations, and since the Menzies government had indemnified the British Government at the outset against claims by Australians, the present Australian Government is now liable for compensation claims.

Jeffrey Sellar

Japanese research

Call for industrial university

Tokyo

CONCERN that Japan's international competitiveness may soon suffer because of a shortage of trained researchers — particularly in biotechnology — has led the Ministry of International Trade and Industry (MITI) to back a proposal that it should set up its own "high-technology" university partnership with private industry.

The proposal comes in a report commissioned by the National Institute of Research Advancement (NIRA), a semi-private body that advises the prime minister's Economic Planning Agency. MITI hopes that action can be taken quickly and that the new university, provisionally called the Institute of High Technology, can begin life as a research institute in Tsukuba Science City before the massive Tsukuba International Science Expo comes to an end in September next year.

Higher education has always been the responsibility of the Ministry of Education, Science and Culture (MESC), and the universities are already awash with unemployed research workers (see *Nature* 21 June, p.659). The problem is, however, that relationships between industry and the universities are generally poor. Very few university laboratories take up joint research with industry (although there are notable exceptions), and there are no Western-style high-technology parks built alongside universities to encourage a close research relationship.

The problem is aggravated by the legal status of academics as "civil servants" employed by MESC, who may not receive payments from nor involve themselves in consulting activities for non-government organizations. Although there have always been ways to avoid direct involvement — inter-university computer centres, for example, received their computers at low prices in return for helping manufacturers

to solve new networking problems — only in the past few years has MESC begun vigorously to encourage university-industry links and to make it easy for those in industry to spend time in the most advanced university laboratories and research institutes.

To complete the picture, industry has often taken a rather dim view of university research. A survey conducted for the 1978 white paper on science and technology showed that more than 60 per cent of industrial respondents felt that university research was of no use to them. A corollary is industry's suspicion (now apparently fading) of holders of doctoral degrees — those with a bachelor's or a master's degree are felt to be much better material for entry into industrial research laboratories.

The head of the team that produced the report, Emeritus Professor Keichi Oshima of Tokyo University, also criticized existing universities for their lack of flexibility in responding to technological innovations. The university professorial "chair" system, with its hierarchical structure, has often been attacked for insulating professors from changes in the outside world.

The new university called for by the report would be supported entirely by private industry and would receive no subsidies — and presumably little influence — from the Ministry of Education, Culture and Science. Its staff will mainly be drawn from the 16 research institutes of MITI's Agency of Industrial Science and Technology which together employ around 2,700 researchers, most of them in brand-new institutes in Tsukuba Science City. Other staff will be provided by private industry. If the new university gets the go-ahead, it will begin life as a research institute, then a graduate school, eventually recruiting direct from high-school students aiming at becoming "top-notch high-technology researchers". Alun Anderson

UK electronics

Sleeping giant waking up?

ONE of Britain's largest technical companies, the General Electric Company Limited (GEC), seems in the process of emerging from two decades of well-managed torpidity. No relation of GE (for General Electric) of the United States, the company has in the past few weeks startled its shareholders by making a bid for British Aerospace, still partly owned by the British Government, and (last week) by saying that it hopes to spend £300 million buying back its own shares from shareholders. But behind the scenes, and in the long run much more important, GEC has set about the rejuvenation of its research programme with unfamiliar zeal.

Not before time, the company's critics will be saying. GEC now is the result of the 1964 Labour Government's widely-shared trust in the management style of Mr Arnold Weinstock (now Lord Weinstock), GEC's managing director then and now, and of its belief that successful companies are preferably big. Thus GEC is an amalgam of its old namesake with Metropolitan Vickers and English Electric (which had already taken over Marconi when swallowed up).

Latterly, GEC's admirers have turned to grumbling. The most widely publicized complaint is that the company's success at making money (profits last year of £671 million on a turnover of £5,600 million) has not been matched by its willingness to invest in new enterprises. One consequence

is that GEC's reserves of cash and negotiable financial assets, commonly called its cash mountain, amount to more than £1,500 million, more than a third of total assets. Shareholders complain that they could do as well by investing their own money in the financial markets. GEC's



Derek Roberts, the new favourite

technical people, however, argue that other large electronics companies (Hitachi and Siemens, for example) have financial reserves which are twice as large when measured by the yardstick of the labour force — more like £15,000 per employee than GEC's £7,000.

For GEC, it is (or should be) more

worrying that its growth has slackened off (both turnover and profits were virtually identical in the past two financial years) and has, over the past five years, been much slower than that of other British electronics companies such as Racal, Ferranti and Standard Telephones and Cables (STC).

Lord Weinstock, meanwhile, is no longer Whitehall's favourite manager — and not simply because of his opposition last year to the Telecommunications Bill, then passing through Parliament, on the grounds that the British telephone network was destined to be an uncontrollable private monopoly. Questions have also been raised about the once-admired management technique, involving the deliberate devolution of financial responsibility to the host of companies in the group. This, the argument goes, may be a recipe for not losing money but inhibits concentration on innovation.

In reality, with a total of 17,000 technically-trained employees, GEC spends something like 11 per cent of its turnover on research and development, much of it on defence contracts. Devolved management has inevitably left individual operating companies with responsibility for the size and character of their own research programmes. While there is no sign that this principle is about to be abandoned, the past two years have seen a marked increase of GEC's centrally supported research.

In the British electronics industry, people say that the arrival of Mr Derek H. Roberts as GEC's director of research five years ago has been the chief catalyst of change. Since then, centrally directed research, previously negligible, has come to occupy a quarter of the effort of GEC's central research organization, chiefly based at the Hirst Research Centre (Wembley, London) and the Marconi Research Centre in Essex.

New efforts are concentrated in micro-electronics, with the development of devices based on gallium arsenide (GaAs) roughly on a par with those based on silicon. Although not a volume manufacturer of microelectronic chips, GEC is an important source of chips designed for specialized applications (many of them in defence). Of the £55 million spent last year by the central research organization, roughly a half was recovered from GEC's operating companies as the cost of commissioned research, with the remainder divided equally between externally commissioned research (from the Ministry of Defence among others) and the company's own new research effort.

Time will no doubt tell how decisively these developments will change the unflattering reputation that GEC has recently required. Last year, electronics and telecommunications accounted for £290 million of the group's profit, and is the most quickly growing sector of its business.

John Maddox

Conglomerate research revives

IF Lord Weinstock is no longer the British Government's favourite manager, Mr Derek Roberts, GEC's director of research and a member of the company's board since the end of last year, is fast becoming the establishment's favourite industrial scientist. A member of the Advisory Board for the Research Councils, in that capacity and in his own right he repeatedly appears as a witness before parliamentary committees, arguing the case for more generous (and rational) policies in the support of science.

Roberts, a vigorous Mancunian physicist, was recruited to GEC from the Plessey Company in 1979 and elected to the Royal Society and the Fellowship of Engineering the following year. Appropriately, he argues for the abolition of the conventional British distinction between science and engineering, as well as for the recognition that women have a potentially important role in industrial research and development. (A brochure describing the successful careers of a score of GEC women has apparently been a hit with British schools.)

The role of a technical director in a conglomerate such as GEC is a little like that of a conductor of an orchestra put

together from musicians whose reputations have been earned as soloists. Roberts reckons to provide searching technical audits of designs emanating from member companies while seeking to sell to them the benefits of research commissions with the central organization.

One of his innovations is a system for encouraging the transfer of technology within the group, partly by means of technical newsletters produced in-house. Intra-group seminars are now commonplace.

Microelectronics is the most obvious theme in Roberts's message to GEC. The Hirst Research Centre at Wembley, architecturally a monument to the 1920s, is being converted internally into a labyrinth of air-conditioned cells for depositing films of this or that on wafers of silicon or other semiconductors. The equipment appears to come in units costing £500,000 a cell, and the transformation is clearly far from complete.

Even so, the technical people are hoping that the British Government will let their company acquire British Aerospace. Unlike the London Stock Exchange, they take the line that in the fields in which they operate, size is still an asset. □

Ageing Japan

Riches make for longevity

Tokyo

THE life expectancy of both Japanese men and women is now the highest in the world, according to figures released by the Ministry of Health and Welfare just a couple of days after the 119th birthday of the world's oldest living person — Shigechiyo Izumi of Kagoshima, on the southern island of Kyushu. Life expectancy is now 78.8 years for women and 74.2 years for men: and if cancer, strokes and heart attacks could be overcome, then Japanese women could expect on average to live to 88.4 years and men to 83.8 years.

But a second report — the white paper on population — from the same ministry's Council of Population Problems, due to be officially released later this month, draws some wider and perhaps more worrying conclusions from current demographic trends. Japan's change in life expectancy has been so rapid — it did not reach 50 years until 1947, and 70 until 1971 — and the postwar decline in population growth so sudden, that the nation is now well on the way to becoming the first "geriatric society". At present, there are 7.5 "workers" (people between 15 and 64) for each person over 65; in the year 2000 there will be only 4.9 and in 2025 2.9 workers for each "old" person.

At present Japan is still a "young" society and several Western nations are already approaching the proportions of old people that Japan will have in 2000. What is so worrying is the speed with which Japan is ageing — some two to four times faster than Western nations — and the consequences this will have for a nation increasingly dependent on high technology and the ability to take up new ideas commonly associated with youth. Already the government is taking the long-term future into account by changing health insurance and pension laws so that it is not left in twenty years with bills it cannot pay. But there are no concrete proposals on how the huge numbers of people in their 60s, 70s and 80s are to play the active part in society the white paper calls for.

The white paper is in part aimed at the United Nations conference on population in Mexico next month. There, it is hoped that Japan's change from a high-birth high-death society to a low-birth low-death society in a very short space of time will be held up as an example to nations with rapidly increasing populations. Japan's population growth rate is now down to 0.7 per cent annually and is still declining. Population should continue to grow from the present 119 million to 130 million by the end of the century and then slowly fall again to the present level. The report also charts, for the first time in Japan, the strong relationship between life expectancy and income levels.

The changes in Japanese life expectancy have, of course, largely been attained by reduction of stillbirths and of deaths during birth and the early years of life, so that the gains in life expectancy for those who survive these dangerous years, compared, say, with somebody living a hundred years ago, are not as great as might first appear. Nevertheless, putting the life expectancy figures together with the average age of marriage given in the report — 25.3 years for women and 28.9 for men — shows that the average couple can look forward to nearly 48 years of married life — surely the most convincing argument there could be for the practice of "serial monogamy".

Alun Anderson

Chinese research administration

Relics of cultural revolution

CHINA'S reinstatement of intellectuals victimized during the Cultural Revolution has sometimes encountered unexpected snags. Although the Chinese press often refers to people trained in university or technical colleges who encounter prejudice among the industrial work-force, odd traces of the Cultural Revolution still remain within the academic establishment itself.

Last month it was revealed that, until May this year the Tianjin Internal Combustion Engine Research Institute was still being run by a "revolutionary committee" whose members had for the most part been appointed during the Cultural Revolution. These people, Beijing radio reported, were "seriously factionalist" in outlook and "despised intellectuals". Talented scientists and technologists found it impossible to carry out their research properly, and were "despised and expelled". There was no proper financial accounting of the institute's budget, and much valuable equipment was carelessly destroyed.

This "state of chaos", the commentator explained, had been allowed to continue because of a bureaucratic muddle about responsibility. Originally, the institute had been under the joint jurisdiction of Tianjin University and the Tianjin Municipal First Machine-Building Industry Bureau, neither of which "attended to its task". In 1978, the Ministry of Machine Building decided unilaterally to take over the institute, but the Ministry of Education objected.

In the following year, Tianjin municipality decided to make it a part of the university, but the Tianjin Municipal Science and Technology Commission objected.

A prolonged period of haggling ensued between the educational establishment and the industrial administrators, neither being willing to let the institute go to the other. In

Polish science

Administration changes afoot

POLAND is likely to acquire two new high-level bodies dealing with the administration of science — a State Committee for Science and Technical Progress and a Council for Fundamental Research. New legislation on science and on the role of the Polish Academy of Sciences is now under way, and although no official announcement has yet been made, an extensive interview with Dr Jan Karol Kostrzewski, the new president of the academy, in the weekly *Polityka* last month gives a strong hint of how the decision will go.

Poland is unique in the Socialist bloc in having no State Committee for Science. Such a committee did, at one time, exist,

February 1982, the state of affairs came to the attention of "leading comrades" in Beijing, who instructed the departments concerned to stop haggling and to sort out the chaos at the institute. Nothing was done, however, and the haggling continued.

At the end of April, Beijing finally intervened. The State Scientific and Technological Commission, the Ministry of Education, the Ministry of the Machine-Building Industry and the Tianjin Municipal Party Committee were ordered to start direct consultations and given a deadline for solving the problem.

On 5 May, it was announced that the institute had been transferred to the Ministry of Education and would be "managed" by Tianjin University.

Finally, on 8 May, the university sent in "capable cadres" to run the institute and, on 17 May, the new administration of the institute and the branch Party committee were announced.

The institute will now be run by a five-member council, four members of which are academics. The new director of the institute is Shi Shaoxi, president of Tianjin University and a member of the Learned Council of the Chinese Academy of Science.

Although the "18 years of chaos" in the institute has now been resolved, there are clearly high-level fears that this is not an isolated case. The radio reports of the affair included an appeal to all Chinese newspapers to publish the story prominently on their front pages. No reason for this request was given, but in a Chinese context it suggests that such reports are meant to prick the consciences of other bureaucrats who might have neglected the institutes under their control. For how many more years these relics will persist is anybody's guess.

Vera Rich

but it was abolished in 1972 and its functions were merged in the new Ministry of Science, Higher Education and Technology. There is no doubt, Kostrzewski told *Polityka*, that the establishment of this ministry as the body which would administer science was a mistake. The proposal to set up the two new bodies is a practical consequence of this conclusion.

The role of the State Committee is fairly obvious. In Kostrzewski's words, "only the state can undertake the burden of financing science and can support in a compact organizational framework the strategic reorientation of our technological policy". The Council for Fundamental Research, however, is a specifically Polish concept, and is, said Kostrzewski, the idea of the academy. Basic research, he explained, has its own specific features.

The new bodies, however, will not apparently mean the disbanding of the Ministry of Science, Higher Education and Technology. In 1981, when the question of reorganizing Polish science was seriously mooted for the first time since the 1972 reforms, there was strong support for the idea that the ministry should be divided, with higher education amalgamated with the existing Ministry of Education and Upbringing, to give a single education ministry covering the whole learning process from kindergarten to *Doctor*

habilitatus. Dr Kostrzewski, however, when asked if the new Council for Fundamental Research might not conflict with the interests of the Ministry of Science, Higher Education and Technology, replied only that the minister, Dr Benon Miskiewicz, was in favour of the new council, and that if the two bodies did have any conflicts, they would "solve them jointly".

The creation of new administrative bodies, however, cannot solve the main problem facing science in Poland — the lack of funds, and, in particular, of hard currency for foreign journals and equipment. In real terms, the science budget for 1982 was only 54 per cent of that for 1978. During the same period, employment in research and development establishments has fallen by 25 per cent, scientists' pay by 37 per cent and that of professors by as much as 60 per cent. Years of underinvestment in industry have led to a constant decline in the country's technological level. According to Dr Kostrzewski, the new "steering mechanism" for science, embodied in the proposed new bodies, offers a chance of remedying the situation. Without a major injection of funds, especially hard currency, however, the new bodies may be of as little use, as one pessimistic academician put it, "as stirring the tea without adding sugar". **Vera Rich**

UK agriculture

Shake-up for research

A RADICAL transformation in the direction of British agricultural research is promised by the announcement, earlier this week, of the formation of a new government committee, called the Priorities Board for Research and Development in Agriculture and Food. The new committee, whose chairman will be Mr Kenneth Durham, chairman of Unilever, will have the unusual task of giving advice both to the three British ministries of agriculture (one each for Scotland and Northern Ireland and one for the rest of the United Kingdom) and the Agricultural and Food Research Council (AFRC).

The creation of the new board seems to have been brought about by the long-standing difficulties in the relationship between the agriculture ministries and the research council. Attempts in the past decade to follow the Rothschild principle that much of the research council's activity would be research commissioned by the ministries have been repeatedly frustrated.

Research planning within the Ministry of Agriculture, Fisheries and Food (MAFF) are said to have been frustrated by the down-grading of the post of chief scientist, while the ministry's own Agricultural Development and Advisory Service (ADAS) has not been as influential as it could have been.

Relationships have been further complicated by the intervention of a body called the Joint Consultative Organisation, originally intended as a means by which farmers could be consulted about research policy, which has become a more than candid critic of AFRC. It is not clear which parts of that organization will continue under the new arrangements.

The membership of the priorities will include Sir Ralph Riley, the retiring secretary of AFRC, and two members of the council, Sir Hans Kornberg (Cambridge) and Mr Ronald Halstead (chairman-designate of Beecham), who will be resigning from the council. The other members are Professor Ronald Bell, director-general of ADAS, Dr Alan Raven, scientific adviser to the Scottish department, and two farmers, Mr John Martin (Cambridgeshire) and Mr John Moffit (Northumberland).

The importance of the new arrangements stems from the way in which the priorities board, by offering advice both to government departments and the research council, will be able to influence the whole pattern of agricultural research in Britain. It should be able also to defend the research council from some of its wilder critics, but also to foster the rationalization of ADAS and the research council. Whether it will wring more money from the Treasury is another matter. **John Maddox**

Yugoslav dissent

Physicist's passport confiscated

DR Ivan Supek, the most senior member of the Yugoslav Academy of Sciences and founder of the Pugwash movement in Yugoslavia, has been deprived of his passport and called in for police interrogation, as part of a new clampdown on intellectual dissent.

Dr Supek, originally a physicist, is one of the earliest known campaigners against nuclear arms. During the Second World War, he was instrumental in passing on to the Western Allies a warning from Heisenberg that Nazi Germany was working on a fission bomb.

In 1958, Supek resigned from the Yugoslav Atomic Energy Commission, fearing that even peaceful nuclear research could too easily be diverted to military ends, and switched his attention to the history and philosophy of science, founding (in 1962) the Institute for Post-graduate Studies at Dubrovnik, whose courses have a strong orientation towards the ethics of science. During the past few years, he has become increasingly critical of contemporary Yugoslav politics, and has published two books abroad, *A heretic of the left*, and a "documentary novel" about the pre-war Croat communist leader Andrija Hebrang.

The clamp-down against intellectual dissent began on 20 April with a police raid on an unofficial seminar in Belgrade. Twenty-eight people, including the lecturer and dis-

sident Marxist philosopher Milovan Djilas, were taken in for questioning. Six of them, including three students, now face conspiracy charges, and the trial for anti-state activities of a seventh, Dr Vojislav Seselj, a former sociology lecturer at the University of Sarajevo, who is being treated as the main ideologue of the group, opened last week.

The Belgrade seminar is accused of "great Serbian nationalism", an ideology unlikely to inspire sympathy among Croats. Nevertheless, a group of Croat prisoners in the Lepoglava prison protested in May to the then Yugoslav head of state, Mika Spiljak (the office has since rotated), demanding political prisoner status for themselves and expressing their sympathy with the Belgrade group. When this action was punished by solitary confinement, they smuggled out a further appeal to the Secretary-General of the United Nations, again pledging support to the senior participants.

This Croat support for the Belgrade dissidents has apparently provoked the security authorities into action against Croat dissidents still at liberty, including Dr Franjo Tudjman, a historian and former partisan general who had been conditionally released from a three-year prison sentence on health grounds but who has been rearrested, and is now in prison.

Vera Rich

Indo-Soviet yoga

Space exercises still up aloft

THE joint Soviet-Indian space flight last April has aroused considerable interest in the Soviet Union in the yoga exercises performed by cosmonaut Rakesh Sharma. However, Soviet citizens who wish to learn more about yoga will be disappointed. Interviewed last month on the Moscow television programme "Problems, research and solutions", the chairman of the Committee for Physical Culture, Marat Gramov, stressed that yoga would not be introduced in the Soviet Union since the Soviets have their own system of physical culture and because yoga is not only a system of physical education but also "a philosophy which does not tally with our outlook".

Even so, official accounts of the joint flight stress the importance of the yoga experiment for the training of cosmonauts, not only for the study of the reactions of the human body to orbital flight but also as a means of minimizing the stress on the cardiomyoskeletal system, particularly during the most sensitive period in the first seven days of flight.

Long-stay missions aboard orbital space stations have from the beginning been a key element in the Soviet space programme.

The current record, 211 days, was set up by Valentin Lebedev and Anatolii Berezovoi in 1982. There has been a vigorous drive to counteract the effects of long-term weightlessness. Symptoms such as disorientation and mild hallucinations or hyperactivity of the kidneys have been carefully monitored. Prophylactic measures range from the development of special constrictive suits and electric stimulation units to maintain muscle tone to infusions of *Eleutherococcus* (Siberian ginseng) in the diet of trainee cosmonauts.

Non-Soviet partners in the joint Interkosmos flights have made their own contribution to these studies. On the joint Soviet-Polish mission of 1978, the Polish participants carried out an investigation of changes in the sense of taste under conditions of weightlessness. The aborted Soviet-Bulgarian flight of April 1979 meant the shelving of a package of five Bulgarian medical experiments which have not yet been performed. Sharma's yoga exercises were planned as an experiment, with his physiological reactions after exercise compared with those of his fellow crew members.

According to the official Soviet presentation to Unispace-82, the Second United Nations Conference on the Exploration and Peaceful Uses of Outer Space, an important feature of Soviet space medicine is its contribution to general health care and preventive medicine for the population at large.

Vera Rich

UK biotechnology

What market, what share?

THE fear that national strategies for the development of biotechnology will lead to a concentration of attention on a few obvious targets, with the result that "no-one will obtain a reasonable return on his investment", is one of the points made by the authors of a report *Biotechnology and British Industry*, published last week by the UK Science and Engineering Research Council (SERC) (£10). Unusually, the report is the result of a commission by the council's biotechnology directorate from Dr Peter Dunnill of University College London, and Mr Martin Rudd, a professional economist. According to the authors, their most daunting task has been to assemble data about the pattern of manufacturing industry and agriculture in which biotechnology may eventually find application.

The underlying tone of the report is one of bravely suppressed gloom. The objective has been to survey the industrial sectors in which biotechnology has potential, and to interest the industries concerned in their exploitation. Part of the problem is that there are few sectors of British industry where the products manufactured by conventional techniques enjoy such a large share of the international market that biotechnologists will be unambiguously guided to them.

The report nevertheless argues for continuing central support for "bioindustry" on the grounds that many opportunities have arisen in industries (such as food processing) unused to supporting research in the strict sense while, the argument continues, government control of the food, pharmaceutical and waste treatment industries is so direct that it can have a "con-

siderable influence" on commercial development.

The report commends the association of small venture companies with larger traditional manufacturers in the development of new products on the grounds that it may then be possible for the several potential outcomes of a line of development to be exploited by the most appropriate of existing manufacturers.

The report has been able to list close on 50 British newly-formed companies with interests in biotechnology, a dozen of them with academic connections. It acknowledges, however, a shortage of suppliers of equipment, especially in fields such as pilot-scale fermentation plants and related equipment. Although the report was completed before the appearance in the United States earlier this year of the Office of Technology Assessment report on the international biotechnology industry (see *Nature* 307, 402; 1984), Dr Dunnill said earlier this week that he considered that assessment to have belittled British capacity for the supply of reagents for biotechnology.

His own opinion (not included in the report) is that British biotechnology is hamstrung by the conviction of traditional manufacturers that only first-degree graduates make suitable recruits, with the result that academic research and graduate teaching in chemical and biochemical engineering was at a low ebb and the output of trained people on the decline. Dunnill said he hoped to persuade SERC's biotechnology directorate to redress the balance by providing high-level studentships for graduate studies at a meeting later this week. □

Nature index of biotechnology stocks

12-Month high	12-Month low	Company	Close previous month	Close 29 June	Change
14	8	Biogen (Switzerland)	9	8½	-½
2	1⅞	Bio-Logicals (Canada)	1½	1⅞	+⅛
14⅞	8	Bio-Response (USA)	8⅞	8¾	-⅛
14	10	Cetus (USA)	10¼	11⅞	+1⅞
10⅞	4¼	Collaborative Research (USA)	5⅞	5⅞	+½
19⅞	11½	Damon (USA)	11⅞	14¼	+2⅞
26¼	12½	Enzo-Biochem (USA)	14	15	+1
10⅞	5¼	Flow General (USA)	6⅞	6½	+⅞
42¼	28¾	Genentech (USA)	29½	34½	+5
10¾	5½	Genetic Systems (USA)	5¾	5¾	0
17¼	8¼	Genex (USA)	9⅞	9	-⅞
23	12¼	Hybritech (USA)	12⅞	14	+1⅞
16¼	8¾	Molecular Genetics (USA)	9½	10¼	+¾
15½	10	Monoclonal Antibodies (USA)	10¼	12	+1¾
60⅞	42⅞	Novo Industri A/S (Denmark)	43½	44¾	+1¼
22¾	14½	Pharmacia (Sweden)	14½	15⅞	+1⅞

Closing prices are for the last Friday of the month. For over-the-counter stocks, bid price is quoted; for stocks on the American and New York exchanges, the transaction price. *Nature's* weighted index of biotechnology stocks stood at 145 on 29 June, compared with 135 a month earlier. Data from E.F. Hutton, Inc.

Nuclear building-down

SIR — In your 5 April 1984 issue (p.490), Stephen Salter presents an intriguing formula for multilateral nuclear disarmament. The Salter proposal is original and well thought out but is, unfortunately, flawed by the unrealistic assumption that either superpower would allow his strategic nuclear force to be restructured by his adversary. The scheme calls for a level of bilateral collaboration we find highly implausible.

There is another possible path to nuclear arms reduction that we think has not so far received sufficient attention. This path entails the simultaneous build-up of defensive systems and the build-down of offensive weaponry — something we have termed, in a recent issue of *Foreign Policy*, defence-protected build-down (DPB). In other words, were either the United States or the Soviet Union unilaterally and incrementally to reduce offensive capabilities while simultaneously deploying defensive systems, a winding down of the arms race could result without producing instability or vulnerability in the strategic balance.

Arms controllers have long considered the build-up of defensive systems to be destabilizing and prone to exacerbating the arms race. The deployment of defensive systems on a gradual or incremental basis need not be destabilizing, however, if it is accompanied by a concomitant and compensatory arms reduction. Since the initiator would gain no strategic advantage from this strategy, his adversary would have no compelling reason to escalate the arms race. We would simply be exchanging one kind of parity (based on mutually assured destruction) for another (based on mutually assured survival).

In a simplified example of how DPB might work, let us assume that the United States and the Soviet Union have achieved parity with 1,000 warheads each. A US-deployed defensive system capable of destroying 10 per cent of the Soviet warheads in an all-out Soviet attack would leave Moscow with only 900 deliverable weapons. This situation would permit Washington to dismantle 100 of its warheads and still maintain the offensive balance. We would not advocate that the United States should continue DPB indefinitely in the absence of a positive Soviet response. Although DPB can be initiated unilaterally, it must eventually evoke a willingness on the part of the other side to join in.

The difficulty in estimating the technical effectiveness of a defensive system is, of course, a serious potential shortcoming of a DPB strategy. Given uncertainties and the prevailing use of worst-case analysis, US planners in the above example will tend to provide low estimates of the defence's technical effectiveness (since it would then dismantle fewer of its warheads than if the effectiveness were high), whereas the

Soviet Union, in contemplating a response to the US action, would naturally estimate the effectiveness to be high. By initially proceeding with DPB in very small increments, however, we can avoid serious discrepancies and stay within the confines of the 1972 Anti-Ballistic Missile (ABM) Treaty, while assessing the Soviet response. The importance of DPB in the beginning resides not in the level of protection provided but in the initiation of a winding-down process and the establishment of a new international norm of behaviour based on offensive arms reduction rather than expansion.

The DPB strategy asserts that defensive systems need not be perfect, or even close to perfect, to achieve a most important goal; namely, a reduction in the level of superpower confrontation and a lessening of the catastrophe that befalls all of us in the event that deterrence fails. The DPB strategy, therefore, is not intended to be a "quick fix" to eliminating our vulnerability. Rather it is seen as a vehicle for making a long-term and evolutionary transition to a safer, defence-oriented world.

JACK N. BARKENBUS
ALVIN M. WEINBERG

*Institute for Energy Analysis,
PO Box 117,
Oak Ridge,
Tennessee 37830, USA*

Nuclear proliferation

SIR — As a result of a misunderstanding, you altered part of my letter on the Non-Proliferation Treaty (NPT) (*Nature* 26 April, p.768), and in the process altered the meaning intended.

In the penultimate paragraph as printed, the implication conveyed is my concern that both Pakistan and Argentina now have the facilities to make nuclear weapons. However, in my original letter I stated that this being the case "it is crucial that the (current) nuclear weapons states show they do not intend to take military advantage of their own civil nuclear programmes, whilst demanding through Articles I and III of the NPT that (current) non-nuclear weapon states forgo the opportunity to do so. The obvious inequity in this situation will continue to keep important non-signatories in opposition to the discriminatory nature of the NPT."

In other words, my concern is as much with the activities of nuclear weapon states signatory to the NPT as with the possible emergent nuclear weapon states.

The importance of this difference in emphasis is that all too often the non-proliferation measures advanced by nuclear weapon states strongly suggest that states which covet nuclear weapons should be regarded as pariahs. My point is that this

excoriation will be treated with the contemptuous rejection it deserves if the current nuclear weapon states continue with the policy best summed up in the epigram: "Our bombs, okay. Your bombs, no way."

DAVID LOWRY

*Energy Research Group,
The Open University,
Walton Hall,
Milton Keynes MK7 6AA, UK*

Homer's sea

SIR — Homer's phrase *oinopa ponton* contains no reference to colour; "wine-dark" is an interpretative English translation (a pretty one, admittedly). It is surely more likely that the phrase meant something like "wine-faced" (*ops*, face), that is "wine-surfaced". The point of similarity was between the moving and frothy surface of the sea and the bubbly surface of wine, when freshly poured into a goblet or cup of metal or pottery and seen by the drinker from above. Homeric men did not drink out of glasses, and anyone who pours out a goblet of wine and quaffs it will see an opaque liquid with "bubbles winking at the brim". It is a sight the drinker may remember when he sees the sea.

JASPER GRIFFIN

*Balliol College,
Oxford OX1 3BJ, UK*

Black as grapes

SIR — In the letter from me which you printed (*Nature* 17 May, p.204) on the meaning of Homer's "wine-dark" sea, you omitted ten words (by a parableptic error) and turned one of my sentences into a *non sequitur*.

The sentence (in the final paragraph) should have read: 'At *Iliad* 18.562 we find μέλας . . . βουτρες, "bunches of black grapes", and Simonides uses βουτρες with οινωπός ("a bunch of black grapes"); similarly, Sophocles (*Oedipus at Colonus* 674-675) applies οινωπ to ivy-berries. You omitted the part of the sentence dealing with Simonides.'

The point is that Homer uses the adjective μέλας ("black") to describe grapes; Simonides uses the same word for grapes with the adjective οινωπός ("wine-coloured"); and Sophocles uses the similar adjective οινωπ of ivy-berries. Thus, in considering the meaning of οινωπ πόντον, we should bear in mind that οινωπ and its kindred words have definite associations with blackness elsewhere.

MARTIN PULBROOK

*Department of Ancient Classics,
St Patrick's College,
Maynooth,
Co. Kildare, Ireland*

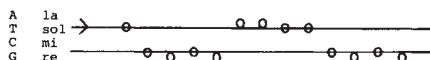
This correspondence is not yet closed — Editor, *Nature*.

Basically musical

SIR — Recent progress in gene cloning and DNA sequencing techniques has produced an enormous amount of base sequence data. In many cases the initial data, read from autoradiograms of sequencing gels, are handwritten or typed on sheets of paper, then transferred to a computer through keyboards. Handling of lengthy and apparently meaningless successions of often thousands of the characters A, T, G and C is tedious and boring. It also requires re-checking to avoid mistakes inherent in such a procedure. We propose an acoustic method to minimize the distress of handling such information.

We have chosen a tone range of a fifth, since this occurs in daily speech. Avoiding a half-tone, three possible arrangements can be considered; bottom-heavy (do, re, mi, sol), symmetrical (re, mi, sol, la) and top-heavy (mi, sol, la, si). Here, we have picked up the second one because of the symmetry to be explored in future use. Then the pitch-to-base assignment we prefer is "re" for G, "mi" for C, "sol" for T and "la" for A.

Using this system, G + C-rich sequences are low-keyed, while thermally less stable A + T-rich sequences are high-keyed. Purine clusters, which tend to distort DNA structure, are leaping and unsettled, while pyrimidine clusters are even and placid. Sequence-melodies can then be transcribed in a two-line score, which is an abridged version of the five-lined score (taking out the upper three lines). For example, the sequence of kinking DNA upon *EcoRI* binding¹, which sounds quite attractive to us, can be written down as follows (the arrow-head on the left indicates 5' to 3' direction).



We can also appreciate the subtlety and delicacy of the variations of consensus sequences like that of the transcription promoters.

One certain advantage of this method is that the sequences are now more easily recognized and memorized. After a few minutes' practice, the longest stretch that we could memorize without difficulty increased at least threefold. In addition, a computer equipped with a sound-generating system can sing back the sequence to facilitate the confirmation by ear. Lastly, this practice may help to bring back some of the pleasure of decoding the mysteries of life from computers.

KENSHI HAYASHI
NOBUO MUNAKATA

Biochemistry Division and
Radiobiology Division,
National Cancer Center
Research Institute,
1-1, Tsukiji 5-chome, Chuo-ku,
Tokyo 104, Japan

1. Frederick, C.A. et al. *Nature* 309, 327 (1984).

All Greek to me

SIR — I was puzzled by the suggestion that the association of the brain's left hemisphere with interpreting phonetic script — Greek alphabetic in particular — may be connected with reading from left to right (*Nature* 309, 409; 1984), when there are both ancient and modern counter-examples such as Aramaic, Hebrew and Arabic, which read from right to left. Second, the ideographs of South-East Asia — the *kanji* characters — which are apparently associated with the right hemisphere, traditionally appear in columns rather than rows, requiring a vertical interpretation. Perhaps more curious are some examples of alphabetic inscriptions surviving from the ancient world, which reverse direction on alternate lines (Southern Arabia, first millennium BC). Examples of this *boustrophedon* style are also found in Ionian Greece. One wonders what this did to the divided brain. More artistically, the Viking *futhark* inscriptions — the runes — are sometimes fantastically entwined, while the Egyptian hieroglyphs were rather versatile, running horizontally or vertically, with pictographs facing left or right. Thus it is perhaps difficult to believe that the brilliantly inventive and literate civilizations of the past (Sumer, Egypt, China) found themselves limited in any sense for lack of an alphabet, whichever way it read.

A.C. SELDEN

*Pound Cottage,
Great Coxwell,
Faringdon,
Oxfordshire, UK*

Abstract policy

SIR — The dialogue among those opposing (M.A. Bray *Nature* 307, 206; 1984), those considering (J.R. Metcalfe 308, 222; 1984), and those approving (S.M. Mould, I. McFarlane 308, 684; 1984) inclusion of abstracts from conference proceedings in comprehensive life science databases, such as the BIOSIS system, has been most informative.

Mould and McFarlane clearly prefer the inclusion of conference abstracts in the database for the values stated by Metcalfe: singular appearance, early notice, awareness of scientific trends. They have no reservations about this practice. In fact, as Ms Mould indicates, the important objective of any such endeavour is to locate and identify *any* published material in a given scientific discipline. We strongly concur in this view.

As you have noted, the ability to discriminate among "full" and "abstract-only" types can be gained by a Boolean search specification "not ABSTRACT". All BIOSIS database entries from the conference literature contain this information.

In our view, it is critical that comprehensive database producers such as BIOSIS should build the most complete

archival record possible. To this end, the BIOSIS database will process approximately 90,000 abstract-only items from the conference literature in 1984. The goal of providing awareness of continuing research and fields of scientific concentration should, in our view, take precedence over the format of the contribution record. The number of abstracts which represent papers from refereed journals covered in the BIOSIS database continues to grow and will not be diminished by our policy on "abstracts-only" literature. All primary journals covered by BIOSIS are refereed sources.

H.E. KENNEDY

*Biosciences Information Service,
2100 Arch Street, Philadelphia,
Pennsylvania 19103-1399, USA*

Peer revue

SIR — Clemens recommends establishing a flourishing domestic population of the spider for biological control of flies (*Nature* 31 May, p.394). It is, however, well known that substantial numbers of the human population are arachnidophobes, especially with respect to large and hirsute members of the group.

If Flanders and Swann¹ are to be believed, establishing such a population of spiders might have the unfortunate and unwanted effect of increasing the fly population by means of self-inflicted fatal injury of human arachnidophobes whilst attempting to remove the hirsute arachnid.

MALCOLM IOSSON

*Department of Microbiology,
University of Reading,
London Road,
Reading RG1 5AQ, UK*

1. Flanders, M. & Swann, D. *The Spider in the Bath*.

Bacterial problems

SIR — Concerning your article (24 May, p.296) regarding Judge Sirica's recent ruling on the use of genetically engineered bacteria, I wish to go on record as indicating that a further detailed consideration of the Lindow experiment led me to file a deposition in the case, which was available and considered by Judge Sirica, that in my opinion there was no reason to block those particular experiments. I do continue to feel — and I believe there is general agreement on this point — that the ecological consequences of experiments of this nature need to be considered carefully now and a consensus needs to be reached on what, if any, regulation is necessary. I fully agree with the conclusion of Mr Stephen Budiansky that it would be extremely unfortunate if Mr Rifkin's tactics, which are, to say the least, unorthodox, led to some valid points being disregarded or not given proper consideration.

PETER H. RAVEN

*Missouri Botanical Garden,
St Louis, Missouri 63166-0299, USA*

CERN comes out again on top

With the discovery of the electroweak bosons ($W^\pm$ and Z^0) in the bag, CERN now announces the discovery of the quark called top. What will come next?

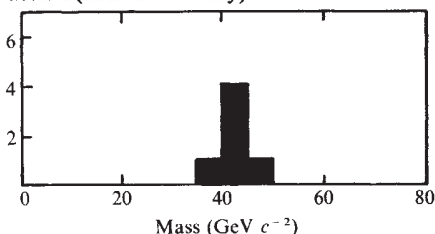
THE Matthew principle — "to him that hath shall be given" — is working in favour of CERN, the European high-energy physics laboratory at Geneva, and of the UA1 collaboration which, at the end of last year, announced the discovery of the $W^\pm$ and Z^0 particles which mediate the electroweak interaction. Last week, the same 80-strong collaboration, under the leadership of Carlo Rubbia, announced the discovery of the missing sixth quark, called *top*, long-predicted but hitherto elusive. By doing so, they have put yet another cap on the electroweak theory while restoring a seemingly symmetry to the evolving picture of quarks as the elementary constituents of the material Universe.

The new development at CERN follows almost exactly along the lines expected (and described, for example, by Dr F. Close in his comment on the electroweak bosons, see *Nature* 303, 656; 1983). The source of the sixth quark is a charged boson, W^+ or W^- , first recognized at CERN by their decay into an electron (with electrical charge of the same sign), with excess momentum carried away by a neutrino. Events of this kind accumulated at CERN in the past two years have amply confirmed that the mass of the $W^\pm$ particles is that predicted by the electroweak theory, the equivalent of 82 ± 2 GeV. The neutral member of the trio of heavy bosons, the Z^0 , is less frequently produced (by a factor of about 10) in the proton-antiproton collisions at CERN, has a greater mass (to the tune of an extra 12 GeV) and is chiefly recognizable by its decay into a pair of electrons, positive and negative.

Although the chief decay path for the $W^\pm$ bosons is that by which their existence was first recognized, it has from the outset been accepted that decay schemes leading to the production of quarks should be recognizable alternatives. Briefly, a W^+ particle should be capable of yielding a *top* and the antimatter version of a *bottom* quark. (W^- would then yield anti-*top* and anti-*bottom*.) For the past two years, there has been general agreement on the way in which these particles could be recognized. The *bottom* quark (or anti-quark) would itself decay into a narrow jet of nuclear matter — pi-mesons for example. And the *top* quark, with a greater mass, would first decay into *bottom* and then yield another jet of particles, this time less tightly collimated. Since the first evidence for $W^\pm$ particles began to accumulate at CERN, people have been wondering whether some

of the events recorded by UA1 were signs of decay of this kind. Six events have now been unambiguously identified as the decay of $W^\pm$ into *top* and *bottom*; the mass of *top*, estimated at 40 GeV, remains substantially uncertain.

For the time being, however, the proof that *top* exists is enough to be going on with. In the simplest terms, the asymmetry that has now been removed is that between the set of known electron-like particles and the set of quarks. For reasons which are frankly not understood, the natural world contains not just one material lepton, the electron (and its anti-particle, the positron), but two others, the muon and the tauon (each with its oppositely charged anti-particle). With each of these three leptons is associated a distinctive neutrino, recognizably different in the mechanisms by which they interact with matter but, on present form, not otherwise distinguishable — they have no electrical charge and no mass. But neutrinos are, like electrons, true leptons — they are involved symmetrically with electrons, muons and tauons in the working of the weak nuclear interaction (as in beta decay).



Distribution of measured top quark mass.

The idea that quarks should also come in pairs, and that there should be as many pairs of quarks as there are pairs of leptons, is more an act of faith than a consequence of theoretical expectation. To be sure, if the world is symmetrical in this way, it is possible to build neater theories, more symmetrical than would otherwise be the case. But that is merely a sign that, in its foundations, theoretical physics remains Pythagorean.

Phenomenologically, the need for symmetry has nevertheless been urgent since the late 1940s. The recognition of the difference between the pi-meson and the muon first raised the puzzle of the apparently superfluous lepton. The discovery (in cosmic rays) at the same time of a new kind of hadronic (nuclear) matter, called *strange* because that is what it was, set the scene for Gillman's radical proposal that mesons such as the pi-meson, but also

the *strange* particles themselves, are pairs of quarks — the pi-meson is a pair called *up* and *down* for example. But nucleons, such as protons and neutrons, and other baryons, are combinations of three quarks — the proton, for example, is two *up* quarks and one *down*. The partner of *strange*, discovered only in 1975, is *charm*. Evidence for *bottom*, also known as *beauty*, was found in 1977 in the proton-proton collisions arranged at Fermilab, where a meson whose mass exceeds the equivalent of 9.4 GeV was surmised to be a bound state of *bottom* and anti-*bottom*.

The quark called *top* (and also, sometimes, *truth*) is thus the missing member of the series. Its appearance has been expected for some time, but is no less welcome to the closet-Pythagoreans on that account. What will, in the short term, matter more is that the steady refinement of the mass now on the cards should make possible a degree of certainty about the nature of some still disputed hadronic particles and resonances. While the electroweak theory itself has been further confirmed, CERN and its UA1 collaboration have provided a more stringent test both of theories of quantum chromodynamics (theories of the strong nuclear interaction) and of Grand Unified Theories (which would roll that together with the electroweak theory but not — yet — with gravitation). Only time will tell whether the outcome is any confirmation of some version or another surprise — yet another pair of leptons or quarks, for example.

Inevitably, the question will arise in Britain whether the discovery of the *top* quark at the collaborative high-energy physics laboratory will bear on the decision, now delegated to a committee under Sir John Kendrew, on whether Britain should continue to collaborate. The arguments run both ways. The discovery of *top* means that CERN's list of unattained achievements has been reduced by one, but at the same time the laboratory's reputation for success has been enhanced. It is, however, unlikely that the committee's recommendations will be determined by scalp-counting of this kind, while high-energy physicists will properly draw attention to the need, now, for the careful understanding of the relationships between the six quarks that will come only from more careful measurements of the decay schemes now recognized, and of the alternatives still to be found.

John Maddox

Gene regulation

Throwing light on plant genes

from Robert Shields

OVER the past year or so there have been several demonstrations that the Ti plasmid of *Agrobacterium tumefaciens*, the bacterium that causes crown gall tumours in plants, can be exploited to insert functional foreign genes into plant cells¹⁻³. The inserted genes are incorporated into the nuclear genome of the plant cells, which, in some cases, have been regenerated into plants and the inserted genes thus transferred by meiosis to the seed⁴. Once the gene is inserted and functioning, the challenge is to determine what features of the gene are necessary for its correct regulation, for example its light response. It does not take a clairvoyant to guess that, as with other eukaryotic genes, the regulation of plant genes is governed by DNA sequences surrounding the gene itself. In two recent papers, one on page 115 of this issue of *Nature*⁵ and one in *Science*⁶, this has been shown to be the case, although a number of intriguing questions remain unanswered.

Both papers deal with the light induction of the genes for the small subunits of ribulose-1,5-bisphosphate carboxylase-oxygenase (Rubisco), the chloroplast enzyme that fixes CO₂ in the photosynthetic cycle. The family of genes for the small subunits is located in the nucleus of the plant cell. Members of the family are regulated by light; individual members are expressed in a tissue-specific manner. Thus, it is found that the amounts of Rubisco are higher in the leaves, pods and peas of the pea plant than in the root and stem⁷.

Brogie *et al.*⁶ describe the insertion of a Rubisco small-subunit gene of the pea (which accounts for about 30 per cent of small-subunit transcripts in pea leaf⁷) into petunia tissue, where the gene becomes light-regulated. There is about 50 times as much specific transcript in illuminated tissue as in dark-held tissue, which is similar to the situation in pea leaf. The inserted gene is transcribed from the normal initiation site; the transcript appears to be correctly spliced and polyadenylated and its protein product is associated with the petunia Rubisco large subunit in the petunia chloroplast. In addition to the coding region, the inserted pea gene has over 1,000 nucleotides of upstream sequence and more than 500 nucleotides of non-coding downstream sequence and it is likely that features responsible for its light regulation lie somewhere in these sequences.

The paper in this issue of *Nature*⁵ describes experiments in which 973 nucleotides lying upstream of a different Rubisco small-subunit gene of the pea (one that is expressed relatively weakly in pea leaves) have been spliced in front of the bacterial gene for chloramphenicol acetyl

transferase; since the presence of chloramphenicol acetyl transferase is easy to measure, its production under the control of the Rubisco upstream sequences can be determined. Sequences downstream of the bacterial gene were provided by the nopaline synthase gene of the Ti plasmid, since these are known to function properly in plants. When the chimaeric gene was inserted via *A. tumefaciens* into the genome of tobacco tissue, the bacterial gene became light-regulated. This is strong evidence that sequences upstream of the Rubisco gene are, at least in part, responsible for its regulation by light.

While Brogie *et al.* show that the extent of regulation of the pea Rubisco genes by light is the same in petunia tissue as in pea leaf, the absolute amount of transcription is far lower. Nevertheless, the plant cell

contains up to five copies of the inserted gene. So although light regulation is independent of tissue type, the absolute amount of transcription may be governed by other factors. What are they? Quite possibly, the gene promoters in the upstream sequences show species-specificity as well as tissue-specificity; pea promoters may just not take to petunia or tobacco. The position of the inserted gene in the foreign genome may also influence the level of transcription. With the availability of disarmed agrobacterium-based vectors that permit the efficient regeneration of plants from genetically transformed tissue, these questions can be approached if not answered. □

1. Herrera-Estrella, L. *et al. Nature* **303**, 209 (1983).
2. Bevan, M.W. *et al. Nature* **304**, 184 (1983).
3. Fraley, R.T. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 4803 (1983).
4. Horsch, R.B. *et al. Science* **223**, 496 (1984).
5. Herrera-Estrella, L. *et al. Nature* **310**, 115 (1984).
6. Brogie, R. *et al. Science* **224**, 838 (1984).
7. Coruzzi, G. *et al. EMBO J.* (in the press).

Robert Shields is at Unilever Research, Colworth House, Sharnbrook, Bedford MK44 1LQ.

Planetary science

Source of the oldest lunar basalt

from S.R. Taylor

THE earliest history of the Moon, as determined from records of craters, basins and uplands established by the Apollo missions, is marked by two very significant events: a period of intense meteoritic or planetesimal bombardment at the close of the formation of the Solar System, and an outpouring of basalt lavas on a global scale. The former preceded the latter, but with a degree of overlap that has been the subject of much discussion. Results reported in a recent paper by L.A. Taylor *et al.* (*Earth planet. Sci. Lett.* **66**, 33; 1983) bear on this issue. Evidence from rocks from the Apollo 14 site in the Fra Mauro region has revealed that some basalt flooding occurred as early as 4.23×10^9 yr ago, far earlier than hitherto believed and well before the end of the bombardment phase.

Younger and more aluminous mare basalts, with ages ranging from 3.85 to 3.96×10^9 yr, had previously been found in the Apollo 14 breccias, indicating that some mare basalt volcanism preceded the last of the large basin-forming collisions. The breccias are located on the Fra Mauro ejecta blanket formed during the impact event responsible for excavating the Imbrium basin at 3.82×10^9 yr.

The close of the great bombardment forms a natural marker in lunar geological history. The next period was dominated by large-scale outpouring of mare basalt lavas, forming the lunar maria which cover 17 per cent of the surface. Some studies have already indicated that mare basalt

volcanism began before the end of the period of the great collisions. For example, basalt sample 10003, the third sample picked up by the Apollo 11 astronauts in Mare Tranquillitatis, gives K-Ar dates of 3.86×10^9 yr, earlier than the date of the Imbrium event. Other evidence comes from the presence of dark halo craters, which were excavated in highland plain units and indicate the presence of buried basaltic lavas. All these observations are consistent with a date of $\sim 4.00 \times 10^9$ yr for the beginning of basaltic volcanic activity on the Moon. The significance of the new results is that they push back the date of eruption of olivine-ilmenite basalts to 4.23×10^9 yr.

The sample analysed bears the typical mare basalt signature of depletion in europium and, by analogy with the Apollo 12 basalts, is derived by partial melting from depths of 200–400 km. The source region of such mare basalts had itself previously undergone a differentiation. The generally accepted explanation is that these pyroxene-olivine-rich regions were formed during the differentiation which followed the large-scale melting of the Moon. This melting, which probably extended to the whole Moon, is conventionally ascribed to accretional melting and is dated at about 4.45×10^9 yr. (This sets the formation of the Moon about 10^8 yr later than the meteorites, in accordance with the Safronov-Wetherill scenario for accretion of the terrestrial planets from a suite of planetesimals.) The subsequent crystalliza-

tion of the Moon resulted in the separation of the feldspathic highland crust, enriched in europium, and the complementary pyroxene-olivine cumulate zone. The latter crystallized at depth to form the source regions, depleted in europium, of the mare basalts. Model Rb-Sr and Sm-Nd ages time the crystallization of these cumulates at about 4.4×10^9 yr.

Various questions arise. The visible extent of mare basalts on the face of the Moon occupies about 1 per cent of the volume of the highland crust, and mare basalt clasts are not abundant in highland breccias from the other landing sites, so how extensive was the early volcanism? What is the energy source? Olivine-rich basalts come from deep within the Moon, well beyond the reach of melting by large-scale impact events. Even the formation of the Procellarum basin, the oldest and largest basin, which forms a framework encompassing much of the nearside topography, is unlikely to have initiated melting at depths of the 200 or more kilometres required by the experimental petrology results. Conventional partial melting induced by radiogenic heat seems a more likely explanation.

A further question of much significance concerns the composition of the highland crust itself. These white feldspathic highlands are comprised of three major components, of which anorthositic feldspar, formed by flotation during crystallization of the magma ocean resulting from the initial whole Moon melting, is dominant. The crust is also strongly enriched in the component known as KREEP, composed of K, U, Th, Ba, Zr, rare earth elements and other incompatible elements. The KREEP component is derived from the last dregs resulting from crystallization of the magma ocean, which pervaded the feldspathic crust, with which they were mixed by the continuing massive meteorite bombardment. The third and most enigmatic component is the so-called Mg suite. It is not directly related to the Mg-Fe phases which crystallized along with the feldspar. While the present bulk highland composition is generally recognized to be due to mixing, the nature of the component responsible for the high Mg content and the primitive Mg/Mg + Fe ratios has remained elusive. Could it be mare basalt? Mixtures of these olivine basalts, perhaps with some admixed dunite, could provide enough Mg to explain the bulk composition of the highland crust. This scenario replaces other explanations for the high Mg content of the highlands, which involve the late addition of planetesimals, or the mixing in of early frozen lunar crusts. Accordingly the discovery of this tiny fragment of an ancient mare basalt may provide fresh insights into lunar crustal history. □

S.R. Taylor is in the Research School of Earth Sciences, The Australian National University, GPO Box 4, Canberra 2601, Australia.

Virology

Epstein-Barr virus in epithelium

from Alan Rickinson

MOST attempts to understand the biology of the interaction between Epstein-Barr virus (EBV) and its hosts begin from a common point of reference, the extreme tropism which this potentially oncogenic agent appears to display for B cells, the antibody-producing lymphocytes¹. Not only is the virus best known for its strong association with Burkitt's lymphoma, a tumour of B-cell origin, but also the life-long virus carrier state, which EBV establishes in all infected individuals, is characterized by a persistent infection of B-lymphoid tissues. To this can be added the evidence, from studies *in vitro*, suggesting that B lymphocytes alone express the EBV receptor and sustain an experimental infection. Significantly, the outcome of the infection is not virus replication but immortalization of the cells into lymphoblastoid cell lines which will carry the EBV genome but which can seldom be induced into a virus-productive cycle. And here's the rub. If EBV replicates so poorly, if at all, in the host cell of its choice, how does it establish a productive infection *in vivo*? For EBV replication clearly does occur in the throat, not just during the course of a primary infection (as witnessed in infectious mononucleosis patients) but also, as recent prospective studies show², very often as a stable accompaniment of the asymptomatic virus-carrier state. Are *in vivo* interactions between EBV and B cells more productive than their *in vitro* counterparts would suggest, or does EBV replication involve a quite different permissive cell type? A recent report from Sixbey and his colleagues at Chapel Hill³ argues strongly in favour of the latter alternative.

Using hybridization *in situ* with specific viral DNA probes, Sixbey *et al.* have been able to detect productive EBV infection in exfoliated epithelial cells in the throat washings of most infectious mononucleosis patients studied, even though this was not always accompanied by demonstrable lymphocyte-immortalizing activity in the cell-free fluid. How the virus might initiate this infection is another question, since no EBV receptors have been detected on human pharyngeal epithelium *in vitro*^{4,5}. One possibility, suggested by the close interaction between lymphoid and epithelial tissues in the pharynx, is that, *in vivo*, the virus is delivered to epithelial cells through fusion with infected B cells. Indeed, one of several Chinese groups active in this field reports that an EBV-producing human B-cell line will spontaneously fuse with nasopharyngeal cells cultured from the tupaia, a lower primate, yielding a hybrid line which continues to produce virus⁶. Despite such

evidence, one feels that if infection of pharyngeal epithelium really is central to the viral strategy, then it should be achieved by a less tortuous route.

Perhaps the most illuminating study to date, again from the Chapel Hill group, shows that primary explants of human cervical epithelium, while completely refractory to 'laboratory strains' of EBV produced *in vitro* by B-cell lines, contains a small proportion of cells which sustain a well documented productive infection following exposure to virus-containing throat washings⁷. Further work is needed to confirm this result, but it would indeed be interesting if the capacity to infect epithelial cells as well as lymphocytes were restricted to natural isolates of EBV, that is to preparations containing virions derived from an epithelial site *in vivo*.

A final dimension to the study of interactions between EBV and epithelial cells stems from the consistent presence of the EBV genome in the malignant epithelial cells of undifferentiated nasopharyngeal carcinoma¹. This is largely unexplored territory; one does not know, for instance, whether the epithelium in which this tumour arises (in the Fossa of Rosenmuller) is a natural site for virus replication or if it only rarely sustains an infection. Nevertheless, the circumstantial evidence to date strongly suggests a role for EBV in the oncogenic sequence.

In an attempt to define that role more precisely, several groups are now transfecting cultured epithelial cells with cloned fragments of the EBV genome and looking for changes in cellular phenotype. As the first results come out, it is clear that we are in for some surprises. Thus, in a recent letter to this journal⁸, Griffin and Karran described the immortalization of monkey kidney epithelial cells by DNA derived from either of two overlapping fragments of the viral genome. Interestingly, the 9 kilobase sequence of overlap, and indeed the whole sequence of one of the fragments, did not include any of the three regions of the EBV genome which are expressed in virus-immortalized B cells (and also, it seems, in nasopharyngeal carcinoma cells) and which code for the putative immortalizing proteins⁹; even those transfected cells which theoretically could express one of the relevant proteins,

1. Epstein, M.A. & Achong, B.G. (eds) *The Epstein-Barr Virus* (Springer, Berlin, 1979).
2. Yao, Q.Y. *et al.* (submitted).
3. Sixbey, J.W. *et al.* *New Engl. J. Med.* **310**, 1225 (1984).
4. Glaser, R. *et al.* *J. natn. Cancer Inst.* **64**, 1085 (1980).
5. Shapiro, I.M. & Volsky, D.J. *Science* **219**, 1225 (1983).
6. Yiwan, L. *et al.* *Scient. sin.* **27**, 284 (1984).
7. Sixbey, J.W. *et al.* *Nature* **306**, 480 (1983).
8. Griffin, B.E. & Karran, L. *Nature* **309**, 78 (1984).
9. Kieff, E. *et al.* *J. infect. Dis.* **146**, 506 (1982).
10. Galloway, D.A. & McDougall, J.K. *Nature* **302**, 21 (1983).

the virus-coded nuclear antigen EBNA, do not appear to do so.

This kind of result, reminiscent of those obtained with rodent cells transfected by fragments of herpes simplex virus DNA¹⁰, raises more problems than it solves. The important issue remains the relevance of this phenomenon to the part which EBV might play in the pathogenesis of nasopharyngeal carcinoma and, as Griffin and

Karran intimate, this should become clearer as their work and that of others extends to include the transfection of human epithelial cells. If not the virus, then at least its students, are losing their B-lymphotropic tag. □

Alan Rickinson is at the Cancer Research Campaign Laboratories, The Medical School, University of Birmingham, Birmingham B15 2TJ.

Molecular biology

New twist to DNA methylation

from Edward E. Max

ON average, the dinucleotide cytidine-guanosine (CpG) is present in vertebrate DNA at a frequency only one-quarter of that predicted by random distribution. This striking 'CpG suppression' was puzzling at first but has since been partially explained¹. A recent analysis² of published vertebrate DNA sequences in our laboratory has added a new puzzle — why do some genes contain localized segments that are CpG-rich regional lapses in the usual CpG suppression?

The partial explanation of CpG suppression involves DNA methylation. Most CpG dinucleotides in vertebrate DNA are methylated as mCpG, in which the mC is 5-methyl cytosine. Because 5-methyl cytosine has a propensity to deaminate to thymidine, CpG dinucleotides in vertebrate DNA tend, over evolutionary time, to be replaced by TpG, the dinucleotide that results from the deamination of mCpG and the complementary CpA (on the other strand). The correlation between methylation, CpG suppression and TpG + CpA elevation in some species was originally pointed out by Bird¹. While the deamination model explains a possible mechanism for CpG suppression in terms of CpG methylation, it leaves open the question of the function of DNA methylation in vertebrates. Many experiments suggest that this function is somehow related to the control of gene expression³.

Our first observation of localized CpG-rich regions within genes came from an analysis of the class I genes of the major histocompatibility complex (MHC). These genes encode highly polymorphic antigens that are major determinants of graft rejection. CpG suppression is absent from all available class I gene nucleotide sequences upstream of the border between exon 3 and intron 3; but normal CpG suppression is found downstream of this border. This is not simply a result of asymmetric frequencies of C or G in these regions; in the *H-2K^d* gene, for example, the ratio of CpGs observed to that expected on the basis of C and G content is about 0.06 downstream of the border between exon 3 and intron 3, but upstream it is 0.96 — that is, essentially no CpG suppression.

Our analysis revealed similar asym-

metries in five mouse, one human and two rabbit class I genes. We also found a CpG-rich region surrounding exon 2 of the genes encoding the β -chain of class II MHC antigens, which are found on B lymphocytes and macrophages and regulate cellular interactions in the immune system. Of the many other genes examined by us and by others⁴, only a few contain CpG-rich regions: the chicken $\alpha 2$ -collagen⁵, hamster adenine phosphoribosyl transferase (Lowy, I. and Axel, R., personal communication) and mouse dihydrofolate reductase^{5,6} genes are three examples.

How do the CpG clusters escape the usual CpG suppression of vertebrate DNA? If CpG suppression requires mCpG as an intermediate, then a simple model to explain CpG clusters would be that these regions of DNA are maintained substantially free of mCpG, that is, are undermethylated, in germ-line DNA. This model predicts that the CpG-rich regions will not show the elevated TpG + CpA levels that result from mCpG deamination, which is precisely what we find. (An alternative model, in which unknown mechanisms generate excess CpG with normal methylation and mCpG deamination, predicts elevated TpG + CpA frequencies.) In the *H-2K^d* gene, for example, the ratio of observed TpG + CpA frequencies to those expected is approximately 1.0 in the CpG-rich 5' region, and about 1.4 in the CpG-suppressed 3' region. Thus, the TpG + CpA frequencies in the CpG-rich regions are consistent with our model of regional germ-line undermethylation.

An experimental test of the proposed model would be to examine directly the methylation state of the CpG-rich regions in germ-line DNA, such as that isolated from sperm. As discussed previously in these columns⁷, methylation of the small fraction of CpG dinucleotides that fall within the recognition sequences of certain restriction enzymes (for example, *HpaII*, *MspI*) can be assessed using Southern blot techniques. The multiplicity of sequences cross-hybridizing to MHC probes complicates such an analysis for MHC genes, but the relevant data already exist for three non-MHC genes that contain

CpG-rich regions. The 5' regions corresponding to CpG-rich segments of the mouse dihydrofolate reductase gene⁸, the hamster adenine phosphoribosyl transferase gene⁸ and the chicken $\alpha 2$ -collagen gene⁹ are, indeed, undermethylated in sperm DNA, whereas the 3' regions are methylated; similar patterns were found for these genes in all other tissues examined. Thus, available data seem consistent with the hypothesis that at least some of the CpG-rich regions are associated with DNA segments that remain undermethylated in the germ-line genome.

Should our model relating CpG-richness and germ-line undermethylation prove correct, it would still represent only the first step towards an explanation of these regions. One would also want to know what features of these regions spare them from the usual methylation of their CpG residues, and what functional significance of the CpG-richness of these regions accounts for their evolutionary preservation. Regarding function, we might first ask whether it is the CpG-richness itself that has critical functional significance, with the undermethylated state perhaps serving only to protect the CpGs from rapid mutation via the mCpG deamination pathway; or conversely, is the undermethylation the critical feature of these regions — perhaps mediating active expression of these genes in sperm cells — with CpG preservation an inconsequential byproduct? Furthermore, is the significance of the CpG-richness the same in MHC and non-MHC genes? It is striking that for both class I and class II MHC genes, the protein domains that exhibit the greatest polymorphisms are the ones encoded in CpG-rich regions; is this more than coincidence?

Given the questions about DNA methylation that remain unanswered in other more thoroughly studied systems, it is evident that additional information about CpG-rich regions and germ-line undermethylation in many genes will be necessary before the full significance of these observations will be understood. Meanwhile, it is worth noting that interesting features can still be gleaned from pure 'sequence gazing' (a pejorative term in some quarters), even from gene sequences as intensively studied as those of the MHC. □

1. Bird, A.P. *Nucleic Acids Res.* **8**, 1499 (1980).
2. Tykocinski, M.L. & Max, E.E. *Nucleic Acids Res.* **12**, 4385 (1984).
3. Bird, A.P. *Nature* **307**, 503 (1984).
4. Smith, T.F., Waterman, M.S. & Sadler, J.R. *Nucleic Acids Res.* **11**, 2205 (1983).
5. Vogeli, G. et al. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5334 (1981).
6. Crouse, G.F., Simonsen, C.C., McEwan, R.N. & Schimke, R.T. *J. biol. Chem.* **257**, 7887 (1982).
7. Felsenfeld, G. & McGhee, J. *Nature* **296**, 602 (1982).
8. Stein, R., Sciaky-Gallili, N., Razin, A. & Cedar, H. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2422 (1983).
9. McKeon, C., Ohkubo, H., Pastan, I. & de Crombrughe, B. *Cell* **29**, 203 (1982).

Edward E. Max is at the National Institute of Allergy and Infectious Diseases, NIH, Bethesda, Maryland 20205.

Vertebrate palaeontology

Rauisuchians and the success of dinosaurs

from Michael Benton

THROUGHOUT most of the Triassic period (245–208 Myr), and before the age of the dinosaurs, the top carnivores were a group of reptiles called rauisuchians. Although this group was formerly thought to have had a restricted distribution, recent papers prove that its distribution was in fact worldwide^{1–8}. More provocatively, some of the new work suggests the need to modify commonly held views about both the evolution of reptiles in the Triassic and the origin of the dinosaurs.

The Rauisuchia were archosaurs (the group of reptiles that also includes dinosaurs and crocodiles) and belonged to the basal group, the 'Thecodontia'. Several of them — such as *Poposaurus*, *Teratosaurus* and numerous unnamed forms that have been identified on the basis of isolated teeth and jaws^{2,3} — were formally classified as dinosaurs. In fact, a whole assemblage of fictitious carnivorous dinosaurs — the Teratosauridae — was founded upon middle and late Triassic rauisuchian teeth and the skeletons of prosauropod dinosaurs^{9,10}. The oldest known true dinosaurs occurred in the middle of the late Triassic (about 225 Myr), and would have been preyed upon by rauisuchians. It has been suggested that rauisuchians were ancestral to (some) dinosaurs¹, but that is unlikely because of their specialized ankles, pelvis and skull. The rauisuchians died out at the end of the Triassic without leaving any descendants, that is as far as we can tell.

Rauisuchians are characterized by several derived characters of the skull (for

example, keyhole-shaped antorbital fenestra, movable joints and extra slot-like fenestra between the premaxilla and maxilla) and pelvis (acetabulum faces ventrally, low ilium, pubis antero-ventral to ilium). There are two families, the quadrupedal Rauisuchidae and the bipedal Poposauridae. A typical late Triassic rauisuchid, *Saurosuchus* (Fig. 1), was 6–7 m long, with a massive skull, short neck, fairly long limbs and a powerful tail.

What is of more importance, from the point of view of archosaur evolution in the Triassic and, in particular, theories about the origin and success of the dinosaurs, is new work showing rauisuchians had an erect gait^{2,3}. The standard view is that advances in the limb posture of archosaurs during the Triassic explain why dinosaurs were so successful. The early Triassic archosaurs had a primitive sprawling posture, like that of living lizards and salamanders. The posture of archosaurs supposedly became semi-erect in the middle Triassic, and finally dinosaurs, with their limbs tucked right under the body, emerged in the Triassic¹¹.

The erect posture is commonly thought to have allowed increased running speed, agility and size, allowing these animals to outcompete all manner of other tetrapods in the middle and late Triassic¹². However, as I have previously argued¹³, large-scale 'competition' of this kind between high-level taxonomic groups and spread over tens of millions of years is most unlikely. More probably, the dinosaurs radiated opportunistically only after the extinction

of earlier dominant groups (mammal-like reptiles, rhynchosaurs, thecodontians). Further weight is given to this view by the fact that rauisuchians had an erect gait and yet were replaced by the dinosaurs.

Dinosaurs and rauisuchians achieved an erect gait in different ways. In a primitive sprawling reptile, the thigh bone (femur) sticks out roughly sideways, and the head of the femur fits straight into the sideways-facing bowl-shaped pelvic socket. In an erect animal, the femur points straight down, and changes have to occur at the joint between femur and pelvis. In dinosaurs (and in mammals), the femur develops a head at right angles to its shaft (Fig. 1a). In rauisuchians, on the other hand, the pelvis is tipped over, so that the pelvic socket faces downwards instead of sideways (Fig. 1b). Recently, it has become evident that erect gait was also achieved by a third group of archosaurs — the saltoposuchid crocodylomorphs, distant relatives of the living crocodiles¹⁴. In the late Triassic, then, three archosaur lineages showed independently derived advances in locomotion, and yet only the dinosaurs survived — the other two lineages died out by the end of the Triassic. □

1. Chatterjee, S. *Nature* **295**, 317 (1982).
2. Bonaparte, J.F. *Rev. Mus. Argent. Cienc. nat. Paleont.* **3**, 55 (1981).
3. Bonaparte, J.F. *J. vert. Paleont.* **3**, 210 (1984).
4. Ochev, V.G. *Paleont. J.* **1982**, 210 (1982).
5. Pinna, G. & Arduini, P. *Natura, Milano* **69**, 73 (1978).
6. Wild, R. *Mesozoic vert. Life* **1**, 15 (1980).
7. Dawley, R.M., Zawiskie, J.M. & Cosgriff, J.W. *J. Paleont.* **53**, 1428 (1979).
8. Dutuit, J.-M. *Annls paléont. vert.* **65**, 55 (1979).
9. Walker, A.D. *Phil. Trans. R. Soc. B248*, 53 (1964).
10. Benton, M.J. *Abstr. 3rd Symp. Mesozoic terrest. Ecosyst.* (in the press).
11. Charig, A.J. in *Studies in Vertebrate Evolution* (eds Joysey, K.A. & Kemp, T.S.) 121 (Oliver & Boyd, Edinburgh, 1972).
12. Charig, A.J. *Mém. Soc. géol. Fr.* **139**, 207 (1980); *Symp. zool. Soc. Lond.* **52** (in the press).
13. Benton, M.J. *Q. Rev. Biol.* **58**, 29 (1983); *Nature* **302**, 16 (1983).
14. Crush, P.J. *Paleontology* **27**, 131 (1984).

Michael Benton is at the University Museum, Parks Road, Oxford OX1 3PW.

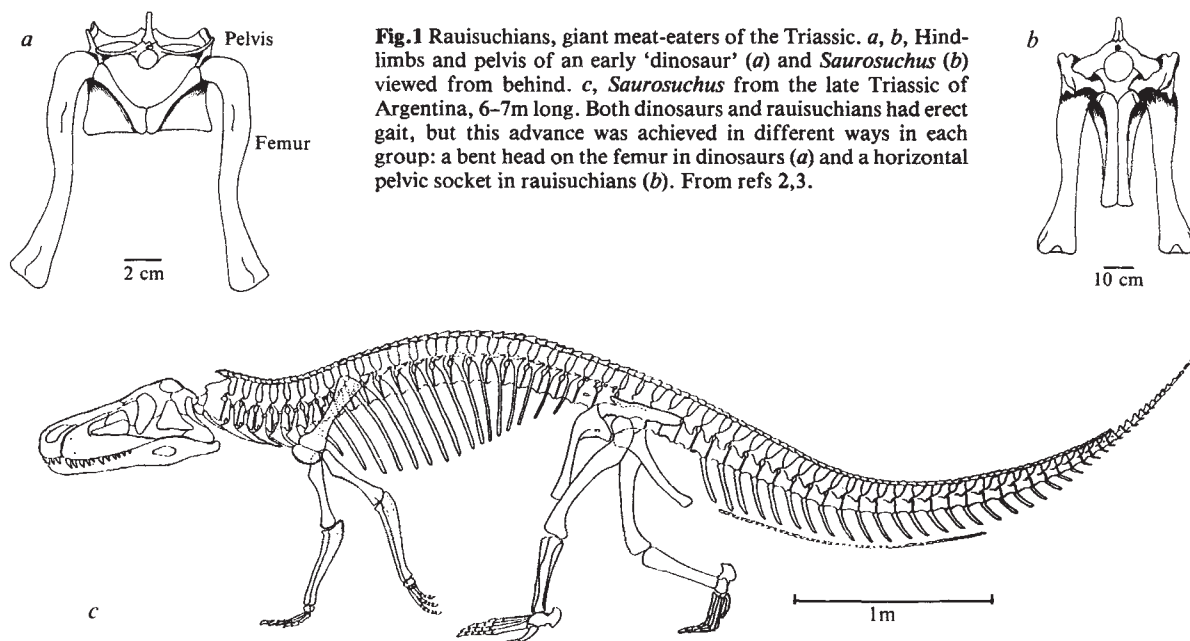


Fig. 1 Rauisuchians, giant meat-eaters of the Triassic. *a, b*, Hindlimbs and pelvis of an early 'dinosaur' (*a*) and *Saurosuchus* (*b*) viewed from behind. *c*, *Saurosuchus* from the late Triassic of Argentina, 6–7 m long. Both dinosaurs and rauisuchians had erect gait, but this advance was achieved in different ways in each group: a bent head on the femur in dinosaurs (*a*) and a horizontal pelvic socket in rauisuchians (*b*). From refs 2,3.

Astronomy

What triggers a quasar?

from C. Martin Gaskell

RECENT observational evidence has greatly strengthened the long-standing view that almost all galaxies have the potential to be quasars (and indeed quite possibly were quasars sometime in the history of the Universe), but that some kind of trigger is needed. The question is, what is the trigger for a galaxy to turn into a quasar or some milder quasar-like object such as a radio galaxy or a Seyfert? A new study by Kennicutt and Keel¹ supports the old idea that interactions between companion galaxies are a very important way of stimulating quasar-like activity at all levels^{2,3}.

The standard model of the tremendous energy release in quasar activity involves accretion of matter onto a massive black hole⁴, with 10^6 – 10^9 the mass of the Sun⁵. Such black holes probably form as the result of stellar collisions in the densely packed galactic nucleus. It has been known for a long time that quasar activity was much more common in the past than it is now^{6,7}. Since massive black holes do not vanish overnight there must be many dormant black holes sitting in galactic nuclei that are not currently showing quasar activity. Our own Galaxy could well have such a dormant black hole since recent spectroscopic surveys of complete samples of galaxies^{8,9} show a low level of nuclear activity in perhaps a half of all spiral galaxies. Therefore, although the existence of a giant black hole is believed to be a necessary condition for quasar activity, it is obviously not a sufficient one. Something else has to happen — the 'monster' has to be fed.

How do you feed a giant black hole? The possibilities considered are too numerous to list, but recent observational studies are reviving the idea that collisions and close interaction between galaxies are responsible. This idea, originally proposed in 1954 by Baade and Minkowski¹⁰ for one of the first radio galaxies identified (Cygnus A), subsequently fell into disfavour¹¹. However, recent high-resolution faint-imaging surveys of relatively nearby quasars show that over 30 per cent of quasars are currently interacting with a nearby galaxy^{3,12}. Seyfert galaxies, too, have close companions¹³; and surveys of samples of nearby galaxies independently suggest that galaxies in close groups and

pairs possess both stronger-than-average central radio sources¹⁴ and stronger-than-average nuclear emission-line activity⁸.

Now, Kennicutt and Keel¹ have made complete spectroscopic surveys of three samples of galaxies: a control sample of non-interacting galaxies; a sample of close pairs; and a sample of closer pairs, showing visible signs of strong tidal interactions. The results are very striking. As expected, about 4 per cent of the control sample show strong Seyfert (quasar-like) activity; but, interestingly, as many as 11 per cent of the close pairs show such activity and a staggering one-third of the tidally-interacting pairs are Seyferts. Not only are there more cases of the most violent nuclear activity in the closer pairs but almost all galaxies in close pairs seem to show enhanced activity, as shown by their emission-line properties. In short, this survey has shown not only that active galaxies are frequently interacting, but also that interacting galaxies are frequently active.

It is now very clear that what happens in the inner few light years of a galactic nucleus is not independent of what happens in the outer edges of the galaxy, a hun-

dred thousand light years away. It seems that tidal disruption of a galactic disc in a close encounter can drive gas into the nuclear regions and either fuel quasar activity or at least cause a burst of star formation^{2,15}. Both theoretical and multi-wavelength observational studies of the details of this fuelling process will be a rich field of research for the next few years.

Meanwhile, lest any aspiring reader be tempted to suggest that a close encounter of our Galaxy with another triggered quasar activity and brought about the extinction of the dinosaurs, let me say that quasar activity has already been shown to be totally harmless to life on Earth^{16,17}. □

1. Kennicutt, R.C. & Keel W.C. *Astrophys. J. Lett.* **279**, L5 (1984).
2. Toomre, A. & Toomre, J. *Astrophys. J.* **178**, 623 (1972).
3. Stockton, A. *Astrophys. J.* **257**, 33 (1982).
4. Salpeter, E.E. *Astrophys. J.* **140**, 769 (1964).
5. Setti, G. & Woltjer, L. *Astrophys. J.* **144**, 838 (1966).
6. Clarke, R.W. & Ryle, M. *Mon. Not. R. astr. Soc.* **122**, 349 (1961).
7. Schmidt, M. *Astrophys. J.* **151**, 393 (1968).
8. Stauffer, J. R. *Astrophys. J.* **262**, 66 (1982).
9. Keel, W.C. *Astrophys. J.* **269**, 466 (1983).
10. Baade, W. & Minkowski, R. *Astrophys. J.* **119**, 206 (1954).
11. Osterbrock, D.E. & Miller, J.S. *Astrophys. J.* **197**, 535 (1975).
12. Hutchings, J.B. & Campbell, B. *Nature* **303**, 584 (1983).
13. Dahari, O. *Astr. J.* **89** (in the press).
14. Hummel, E. *Astr. Astrophys.* **96**, 111 (1981).
15. Shklovsky, I.S. *Soviet Astr.* **6**, 465 (1962).
16. Fairall, A.P. *The Observatory* **100**, 7 (1980).
17. Smith, R.C. *The Observatory* **100**, 123 (1980).

C. Martin Gaskell is in the Department of Astronomy, University of Texas, Austin, Texas 78712, USA.



100 years ago

ON THE EVOLUTION OF FORMS OF ORNAMENT

THE leaf in *Dracunculus* has a very peculiar shape: it consists of a number of lobes which are disposed upon a stalk which is more or less forked (tends more or less to dichotomise). If you call to your minds some of the Pompeian wall decorations you will perceive that similar forms occur there in all possible variations. Stems are regularly seen in decorations that run perpendicularly, surrounded by leaves of this description. Before this, these suggested the ideal of a misunderstood (or very conventional) perspective representation of a circular flower. Now the form also occurs in this fashion, and thus negates the idea of a perspective representation of a closed flower. It is out of this form in combination with the flower-form that the series of patterns was developed which we have become acquainted with in Roman art, especially in the ornament of Titus's Thermae and in the Renaissance period in Raphael's work.

It is difficult to obtain a firm basis on which to conduct our investigations from the historical or geographical point of view into this form of art, which was introduced into the West by Arabico-Moorish culture, and which has since been further developed here. There is only one

method open to us in the determination of the form, which is to pass gradually from the richly developed and strongly differentiated forms to the smaller and simpler ones, even if these latter should have appeared contemporaneously or even later than the former. Here we have again to refer to the fact that has already been mentioned, to wit, that Oriental art remained stationary throughout long periods of time. In point of fact, the simpler forms are invariably characterised by a nearer and nearer approach to the more ancient patterns and also to the natural flower-forms of the Araceae.



From *Nature* **30**, 272, 17 July 1884.

In the article 'Early evolution of leaves' by J. B. Richardson (*Nature* **309**, 749; 1984), there was a mistake in the ninth sentence of the third paragraph. It should read '... and work by Andrews, Gensel, Kasper, Banks and Hueber indicates a high diversity of plants at that time (see *inter alia* ref. 9)'.

Semen polyamines in AIDS pathogenesis

SIR — Recent correspondence has drawn attention to the immunosuppressive properties of cellular and soluble constituents of human semen which may be significant in the aetiology of the acquired immune deficiency syndrome (AIDS)¹⁻³. The discovery of spermatozoa in semen by Antoni van Leeuwenhoek in 1677 was accompanied by a description of small crystals deposited from seminal fluid⁴ which are now known to be composed of spermine phosphate. Spermine is present in seminal plasma at concentrations between 5 and 15 mM together with lower concentrations of another polyamine, spermidine, and their biosynthetic precursor, the diamine putrescine⁵. Micromolar quantities of either polyamine have been shown to inhibit the *in vitro* responses of murine lymphocytes assayed by mitogen stimulation, mixed lymphocyte reactions and cytolytic responses⁶. Such inhibitory effects were shown to be reversible but they occurred only in the presence of calf or fetal calf sera.

Polyamine oxidase activity in ruminant sera may form aminoaldehydes and acrolein which are known to be toxic to cells in culture. Inhibitors of spermine oxidase have been shown to abolish the suppression of lipopolysaccharide-mediated mitogenesis of murine lymphocytes produced by fetal calf serum and polyamines but all suppression was not necessarily due to cytotoxicity⁷. Indeed, the DNA synthetic response of human lymphocytes stimulated by phytohaemagglutinin can be inhibited by spermine or spermidine in the presence of fetal calf serum without apparent cytotoxic effects as measured by Trypan blue exclusion (V. Tan and J.D.W., in preparation). Another recent study has shown that the fully acetylated diamines hexamethylenediacetamide and diacetylputrescine can cause inhibition of murine B lymphocyte activation in the absence of ruminant serum⁸. Thus, although their mechanism of action remains unknown, the polyamines and diamine in human semen can be implicated experimentally in immunosuppression.

The postulated sequelae of the exposure to high concentrations of polyamines in human semen could be exacerbated if similar immunosuppressive mechanisms arise from infections commonly found in AIDS patients. Seroepidemiological studies have shown antibodies to human cytomegalovirus (CMV) in 94% of male homosexuals⁹ and 14 out of 15 autopsied AIDS cases in one study had disseminated CMV infections¹⁰. Infection of MRC5 cells with human CMV results in increased polyamine biosynthesis leading to spermine levels 12-fold higher than in uninfected control cells¹¹. These virus-induced metabolic changes can be calculated to produce enhanced accumulation of polyamines sufficient to be in-

hibitory in terms of the lymphocyte function assays described above. However, other *in vivo* immunomodulatory effects mediated by polyamines can be envisaged. It is known that exogenous putrescine can restore 'suppressor' cell activity associated with the regulation of haematopoiesis after polyamine depletion in mice treated with the specific inhibitor DL- α -difluoromethylornithine¹². A similar requirement for polyamines in the modulation of accessory cell activity in the immune system may influence T-suppressor cell activity in the presence of exogenous polyamines. Immunosuppression is a significant feature of human CMV infection and activation of T-suppressor cells has been shown in cytomegalovirus mononucleosis¹³.

Lymphotropic retroviruses have recently been isolated specifically from AIDS patients but it is still not clear if such infections alone can account fully for the immunopathogenesis of the disease¹⁴. If other co-factors are required, various mechanisms of immunosuppression can be proposed which are mediated by components of semen found in the high-risk group. Since human CMV is present in semen, in addition to saliva and urine, such body fluids constitute a potent source of infection. The presence of polyamines in semen and their further production during human CMV infection may be significant aetiological factors in the pathogenesis of AIDS.

J.D. WILLIAMSON

*Division of Virology,
Department of Medical Microbiology,
St Mary's Hospital Medical School,
London W2 1PG, UK*

1. Shearer, G.M. & Rabson, A.S. *Nature* **308**, 230 (1984).
2. Olsen, G.P. & Shields, J.W. *Nature* **309**, 116 (1984).
3. James, K. *Nature* **309**, 117 (1984).
4. van Leeuwenhoek, A. *Phil. Trans. R. Soc.* **12**, 1040 (1678).
5. Williams-Ashman, H.G. & Lockwood, M.D. *Ann. N.Y. Acad. Sci.* **171**, 882 (1970).
6. Byrd, W.J., Jacobs, D.M. & Amoss, M.S. *Nature* **267**, 621 (1977).
7. Labib, R.S. & Tomasi, T.B. *Eur. J. Immun.* **11**, 266 (1981).
8. Ryan, J.L., Bondy, P.K., Gobran, L. & Canellakis, Z.K. *J. Immun.* **132**, 1888 (1984).
9. Drew, W.L. *et al. Lancet* **ii**, 125 (1982).
10. Macher, A.M. *et al. New. Engl. J. Med.* **309**, 1454 (1983).
11. Tysms, A.S. & Williamson, J.D. *Nature* **297**, 690 (1982).
12. Sharkis, S.J. *et al. Blood* **61**, 604 (1983).
13. Carney, W.P. *et al. J. Immun.* **126**, 2114 (1981).
14. *Lancet* **i**, 1053 (1984).

SIR — I would like to acknowledge an oversight in our recent letter (*Nature* **308**, 230; 1984), in which we did not, but should have cited Mavligit *et al.* (*J. Am. med. Ass.* **251**, 237; 1984) for pointing out the difference between the epithelial lining of the vagina and the rectum.

GENE M. SHEARER
A. RABSON

*National Institutes of Health,
National Cancer Institute,
Bethesda, Maryland 20205, USA*

Caution on AIDS viruses

SIR — Since human T-cell leukaemia virus (HTLV) 1 and HTLV 2 have been isolated from patients with acquired immune deficiency syndrome (AIDS), there are now at least three or possibly four T lymphotropic viruses that can be associated with this syndrome (if lymphadenopathy-associated virus proves to be different from HTLV 3).

If at least three T-cell tropic viruses are recoverable from AIDS patients, albeit with different frequencies, it suggests that they are more likely to represent opportunistic infections or reactivations from latency. This is not to say that they play no role in the overall pathogenesis of AIDS.

The problem of so many isolates is not mitigated by a possible relationship between them, nor by differences in isolation frequency. This difficulty does not seem to bother Dr Groopman (*Nature* **308**, 769; 1984), who appears satisfied that the causation of AIDS has been revealed.

AIDS is indeed a tragic disease, as Dr Groopman notes. This tragedy can only be compounded by premature interpretations of the role of new viral isolates with the promise of a vaccine, and a diminished focus on other possible aetiological factors.

J.A. SONNABEND

*49 West 12th Street,
New York,
New York 10011, USA*

Somatic gene mutation and breast carcinoma

SIR — In her crisp summary on the new information concerning recessive mutations in Wilm's tumour and retinoblastoma and her presentation of likely mechanisms of their evolution, Ellen Solomon (*Nature* **309**, 111; 1984) suggests the possibility that similar mechanisms may lie behind the more common carcinomas of late onset and these should now be explored. She points out that any such efforts are hampered by the 'haystack' problem because of our lack of knowledge about which chromosome to study with DNA probes. The challenge of finding out if the tumour cells contain increased hemi- or homozygosity is reasonable when genes are pinpointed by the breakpoint of a specific chromosome change but, as Ellen Solomon stresses, it is often difficult to obtain adequate preparations from carcinomas so it is not yet known if they carry specific chromosome abnormalities.

We have recently tried to overcome these problems and fill the gap in our knowledge of breast carcinoma by cytogenetic investigations using consecutive banding techniques on direct chromosome preparations from 111 breast carcinomas obtained at mastectomy from 110 women and 1 man (Rodgers *et al. Cancer Genet. Cytogenet.*, in the press).

All tumour samples which contained cells in mitosis (52/111, 47%) showed chromosome abnormalities which in a high

proportion of cases were extremely complex. Nevertheless we were interested to find that among the tumours in the diploid range (20/52, 38%) there were some which could be investigated in detail and a few with only little chromosome change.

Even more interesting was the finding that of the three tumours which carried the least chromosome change the simplest was pseudodiploid, having only monosomy 16 and the marker $\text{del}(1)(\text{qter} \rightarrow \text{p21})$ and identical changes were seen in a second tumour although there was an additional marker in this case. The third tumour was slightly more complex, and included in the seven markers were a chromosome 1 marker $\text{t}(1;2)(\text{qter} \rightarrow \text{1q21}::\text{2q33} \rightarrow \text{2qter})$ and also a deleted chromosome 16, $\text{del}(16)(\text{qter} \rightarrow \text{p21})$.

It seems reasonable to conclude that any search for somatic gene mutations in breast carcinoma should be directed to either chromosome 16 or to the long arm of chromosome 1. As markers involving the long arm of chromosome 1 have also been shown for other carcinomas and malignancies such as myelomas, lymphomas and melanomas, it seems that monosomy and thus hemizyosity for chromosome 16 might be the more relevant.

MAJ HULTEN

Regional Cytogenetics Laboratory,
East Birmingham Hospital,
Bordesley Green East,
Birmingham B9 5ST, UK

Mutualism disagreement

SIR — We have commented previously on May's views about the global distribution of mutualism^{1,2}. In his recent article³, he contrasts "tight mutualistic associations between clearly distinct species" and "entities like lichens, where algae and fungi are fused into something akin to a single organism". He then excludes the latter types from his considerations. This evades the fact that the participants of lichens and similar associations (Table 1 of ref. 4 lists 25 such kinds) still have separate identities and genetics, and distinct (yet interesting) population dynamics and patterns of evolution. Sounder reasons for selection of particular sub-sets from the range of mutualisms available are required.

The supposed (but disputed²) inconspicuousness of ecologically obligate mutualisms in temperate regions is used to provide backing to a theory, also mentioned by May³, which predicts that such mutualisms are dynamically fragile and therefore more likely to be found in tropical environments, assumed to be more stable than temperate ones. Clearly, if this theory is tested on data from which obligate mutualisms common in temperate environments have been excluded, the conclusions reached are unlikely to reflect reality. It would be more instructive to assess those features of the theory which are responsible for such an unrealistic prediction about the geographical distribution of obligate mutualists. One likely

feature is the theory's lack of a spatial dimension, for it can be demonstrated that obligate mutualists are more resilient against environmental changes if they are able to diffuse through space^{5,6}.

R. LAW

Department of Biology,
University of York,
York YO1 5DD, UK

V. HUTSON

Department of Applied Mathematics,
University of Sheffield

D.H. LEWIS

Department of Botany,
University of Sheffield,
Sheffield S10 2TN, UK

1. May, R.M. *Nature* 296, 803 (1982).
2. Lewis, D.H. *Nature* 297, 176 (1982).
3. May, R.M. *Nature* 307, 410 (1984).
4. Law, R. & Lewis, D.H. *Biol. J. Linn. Soc.* 20, 249 (1983).
5. Hutson, V. *SIAM J. appl. Math.* (submitted).
6. Hutson, V., Law, R. & Lewis, D. *Am. Nat.* (submitted).

Archaeopteryx's morphology

SIR — Hecht and Tarsitano¹, in their response to my necessarily brief News and Views article², accuse me of "distorting and misrepresenting (their) viewpoints" concerning the morphology and homologies of the manus of *Archaeopteryx*. In my article I was concerned solely to highlight aspects of disputed morphology which could be settled by a comparison of the actual specimens, if they could be gathered together for the forthcoming Eichstatt symposium.

Nowhere in the article did I refer to the homologies of the digits of *Archaeopteryx*, either with modern birds or theropod dinosaurs; nor did I suggest that *Archaeopteryx* might have a digital reduction pattern of 1.2.3 as Hecht and Tarsitano state¹. I can only assume that because I favoured a phalangeal formula of 2.3.4, Hecht and Tarsitano presumed that I must also favour the preservation of the ancestral phalangeal formula and hence a 1.2.3 digital reduction pattern. However, a phalangeal formula of 2.3.4 can be derived via a 1.2.3 or 2.3.4 digital reduction pattern from a primitive reptilian manus of five digits with a phalangeal formula of 2.3.4.5.3. If the 2.3.4 digital reduction pattern is accepted, then a 2.3.4 phalangeal formula is less derived than would be the case if the broken phalanx hypothesis is accepted, as this would give a phalangeal formula of 2.3.3.

On the question of this supposed break in the topographic third digit, Hecht and Tarsitano are quite definite. "We regard the articulation between the first and second phalanges as displacement and breakage"³ (p.161), and their illustrations (pp.156-157) label the joint as a "break within (the) first phalanx of the fourth digit". After criticism of this point⁴ they become more circumspect, but still argue the case for a break⁵, citing Heller's⁶ mistaken restoration (see plate 7 of ref. 6) of the left manus of the Maxberg specimen as an example of an unbroken phalanx. However, a perusal of the accompanying photographs (plates 9 and 11 of ref. 6)

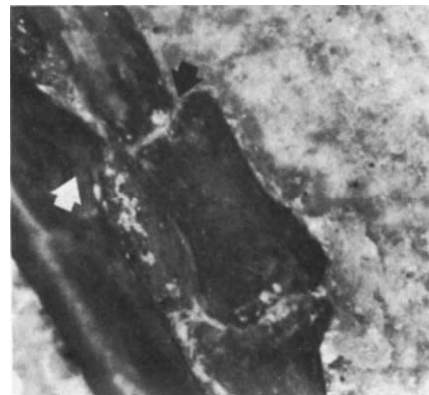


Fig.1 The left manus of the Berlin specimen showing the supposed break (actually a joint, black arrow) in the first phalanx of the third topographic digit and the outgrowth (the "flange" of Hecht and Tarsitano, white arrow) on the first phalanx of the second topographic digit.

reveals the "unbroken phalanx" as two distinct impressions, and my own detailed photographs of the critical area of the Berlin specimen² (Fig.1) show that the supposed break is as good a joint as the other, undisputed joints illustrated.

My second "distortion and misinterpretation" concerns the function of the supposed flange. Hecht and Tarsitano³ regard this structure as "a brace or point of attachment" (p.162). In the absence of any explanation as to what this means I think that I am entitled to describe the function as "to brace and strengthen"². I might add that to brace a part of the wing serves to brace the structure as a whole.

The flange, however, appears to be a pathological outgrowth as it has a different texture from the smooth bone of the phalanx, being more rugose. It also has a distinct margin, which cuts across (overgrows?) a linear structure near the left-hand margin of the exposed bone (see Fig.1).

As for Hecht and Tarsitano's "crash-dive" scenario^{1,5}, a joint is not "caused by forces during stalling (or) . . . impact"¹, only a break. The crossing of the fingers, which can be seen in the Berlin and Eichstatt specimens (both hands) and the Maxberg specimen (one hand), are attributed to the same "catastrophic" cause. This twisting has a much more parsimonious explanation as being due to the heeling over of the strongly arched claws sideways⁴ as the dead animal settled on the sea bed. Hecht and Tarsitano do their case for the homology of avian digits with those of *Archaeopteryx* no good by a myopic view of palaeontology and an extravagant use of *ad hoc* scenarios.

M. E. HOWGATE

Department of Zoology,
University College London,
Gower Street, London WC1 6BT, UK

1. Hecht, M.K. & Tarsitano, S. *Nature* 309, 588 (1984).
2. Howgate, M.E. *Nature* 306, 644 (1983).
3. Tarsitano, S. & Hecht, M.K. *J. Linn. Soc. Zool.* 69, 149 (1979).
4. Thulborn, R.A. & Hamley, T.L. *Aust. J. Zool.* 30, 611 (1982).
5. Hecht, M.K. & Tarsitano, S. *S. Geogios mem. spec.* 6, 141 (1982).
6. Heller, F. *Erlanger Geol. Abh.* 31, 3 (1959).

A microchemical facility for the analysis and synthesis of genes and proteins

M. Hunkapiller*, S. Kent, M. Caruthers*, W. Dreyer, J. Firca*, C. Giffin*, S. Horvath, T. Hunkapiller, P. Tempst & L. Hood

Division of Biology, California Institute of Technology, Pasadena, California 91125, USA

A series of automated instruments that use state-of-the-art chemical methods has been developed for high-sensitivity protein sequencing, DNA synthesis and peptide synthesis. These instruments have been integrated into a centralized microchemical facility in order to promote their use for the study of a variety of biologically interesting problems. This facility has as one of its major functions the development of new chemistries and instrumentation for the structural analysis and synthesis of genes and proteins.

IN physics, the development of centralized facilities housing complex instruments such as synchrotrons and accelerators has promoted collective scientific efforts by researchers with a variety of backgrounds and interests. These sophisticated facilities and the joint efforts they foster have permitted the analysis of matter and forces with increasing levels of sensitivity and have, accordingly, changed our views of the world and of the matter of which it is composed. In a similar manner, the development of sophisticated instrumentation in biology permits the analysis of very small quantities of protein and the rapid synthesis of oligonucleotides and protein—methodologies which are revolutionizing modern biology.

About 10 years ago at Caltech (California Institute of Technology) we set out to develop a new instrument called the protein microsequenator to decrease the amount of protein required for amino acid sequence analysis. Subsequently, we initiated the development of a series of instruments that allow us to synthesize and analyse genes and proteins using more sensitive or effective chemistries than those generally available elsewhere. We call this collection of instruments the Caltech Microchemical Facility (see Table 1); it now includes several gas-phase protein microsequenators, an automated DNA synthesizer, a peptide synthesizer and a computer with sophisticated software for the analysis of protein and gene sequence data. The microchemical facility performs two important functions—state-of-the-art analysis or synthesis and the development of new chemistries and instrumentation; for example, we are presently developing a highly sensitive mass spectrometer for analysis of amino acid derivatives, and an automated DNA sequencer.

Here we describe the instruments comprising the facility and the techniques used.

Amino acid sequence analysis

Our gas-phase protein microsequenator for the direct sequence analysis of proteins and peptides uses Edman degradation¹, a repetitive reaction in which the α -amino group of the polypeptide chain is coupled to phenyl isothiocyanate in alkaline conditions, whereupon the amino acid can be cleaved from the remainder of the peptide with acid (Fig. 1). The resulting amino acid derivative (an anilinothiazolinone) is converted to the more

stable phenylthiohydantoin (PTH) amino acid, which is analysed by HPLC. The remaining peptide is subjected to further degradation cycles to generate its full amino acid sequence.

Per Edman automated this chemical reaction in 1967 by developing the first protein sequenator². His instrument, called a spinning cup (or liquid-phase) sequenator, is a sophisticated device in which an automated controller directs the transfer of liquid reagents and solvents to and from a reaction vessel containing the peptide. While the original sequenator provided a dramatic increase in sensitivity and speed over previous sequencing methods, it required ~100 nmol of protein for sequence analysis. Although commercial versions of the spinning cup instruments can now operate routinely using but a few nanomoles, many proteins of interest to molecular and chemical biologists are even then only available in much smaller quantities. This led us to modify the Edman technique to yield the gas-phase sequenators we now use^{3,4}.

In these machines, the peptide is contained not on the wall of a spinning glass cup but on the surface of a thin disk of glass filter paper mounted inside a small glass column. Polybrene, a polymeric quaternary ammonium salt that adheres strongly to both glass surfaces and proteins, prevents the protein from being dislodged by solvents flowing through the column. The use of Polybrene, originally described for sequencing peptides in the spinning cup sequenator⁵, eliminates the need to covalently attach very small quantities of protein or peptide to a solid support before placing it in the column. The Polybrene support requires that the Edman reagents, particularly the coupling base and cleavage acid, be delivered as vapours (carried by a stream

Table 1 Caltech Biology Microchemical Instrumentation Facility

Instrument	Function
Protein microsequenator	Primary structural analysis (order of amino acids) in proteins and peptides
EOID mass spectrometer	High-sensitivity analysis of complex mixtures of biochemicals
DNA synthesizer	Automated synthesis of oligonucleotides (genes)
Peptide synthesizer	Automated synthesis of peptides and fragments of proteins
Laboratory computer	Analysis of DNA and protein sequence data

* Present addresses: Applied Biosystems, Inc., Foster City, California 94404, USA (M.H.); Department of Chemistry, University of Colorado, Boulder, Colorado 80309, USA (M.C.); Abbott Laboratories, Inc., Chicago, Illinois 60064, USA (J.F.); Jet Propulsion Laboratory, Pasadena, California 91109, USA (C.G.).

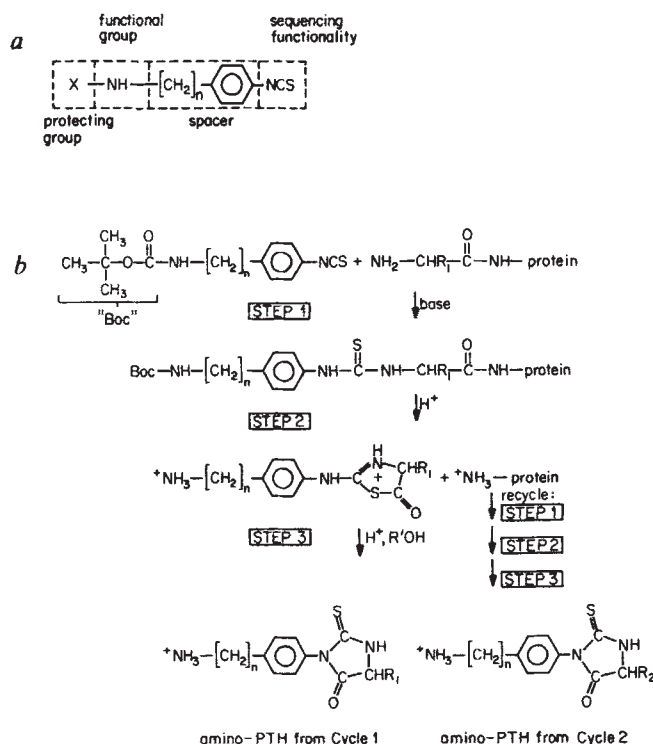


Fig. 1 Modified reagent for improved detectability in protein sequencing. The reagent (**a**) contains a cryptic functional group. In Edman degradation (**b**) the functional group is unblocked, giving rise to 'amino-PTH' derivatives which can be used to yield fluorescent derivatives that are more sensitively detected.

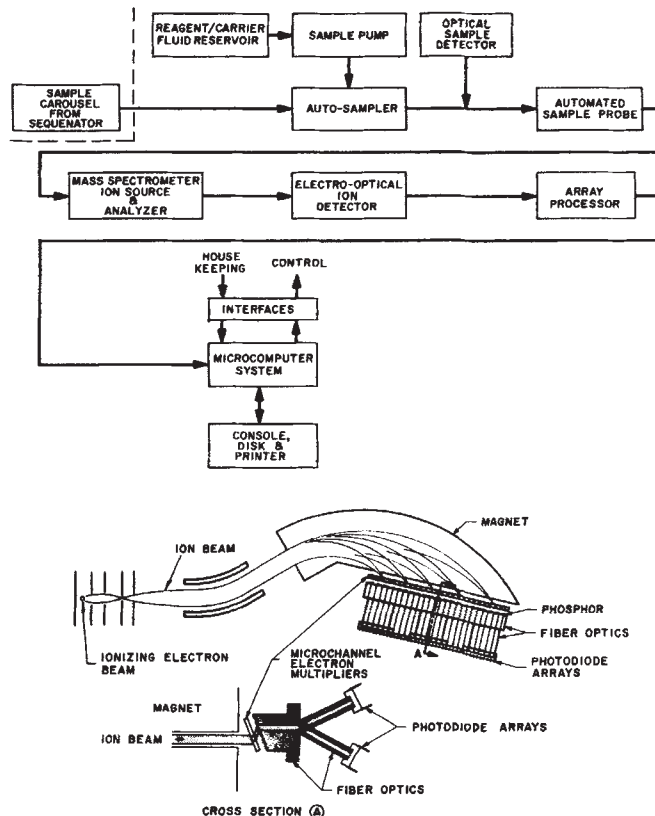


Fig. 2 Electro-optical ion detector array mass spectrometer. A block diagram of the instrument is shown in the upper half of the figure. The lower part shows the geometry of the mass spectrometer and detail of the ion detector. The instrument simultaneously monitors the entire mass range 25–500 AMU and can detect the arrival of a single ion with resolution of 0.1 AMU, giving fmol sensitivity for PTH analysis. The entire instrument (sample handling, acquisition of the mass spectra, data reduction) is computer-operated with a sample turn-around time of less than 5 min.

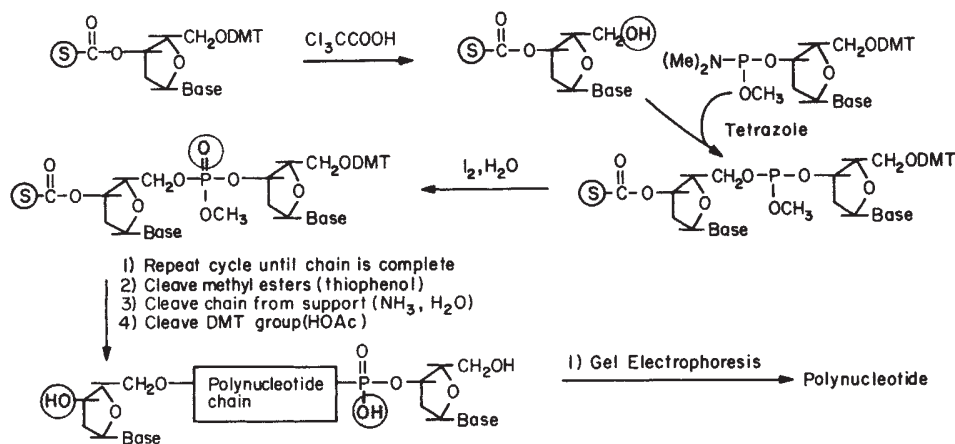
of an inert gas such as argon) which readily diffuse into the Polybrene film containing the peptide. The mechanical flow properties and small size of the column allow quick and efficient solvent flushing after successive steps of the chemical degradation; the anilinothiazolinones produced are carried from the column by a few hundred microlitres of solvent, free of the excess reagents and reaction by-products that contaminate the samples from a spinning cup. This is why subsequent analysis of the PTH amino acids can be performed at the picomole rather than the nanomole level.

The gas-phase sequenator has several advantages over the spinning cup sequenators, the most important of which is sensitivity. The gas-phase instrument requires as little as 5 pmol of most proteins and 20 pmol of many peptides for sequence analysis. Although not all samples yield comparable results, sequence analysis using <100 pmol is now routine. The gas-phase sequenator also exhibits high efficiency as measured by the repetitive cycle yield, which ranges from 98% with 10 nmol of protein to 92% with 10 pmol, resulting in the successful identification of 90 and 21 amino acid residues, respectively, in single analyses³. This ability to determine long stretches of sequences minimizes the need to resort to multiple chemical or enzymatic cleavages to generate many small peptides for further analysis of a protein or large peptide. As a rule of thumb, the sequenator can provide analyses extending to 70 residues or more with a few nanomoles of protein, nearly complete sequences of small or medium-sized peptides (<40 residues) with a few hundred picomoles or less, and analysis of 15–30 residues with as little as 5–20 pmol of proteins and many peptides. A third key feature of the gas-phase sequenator is its versatility. The machine can handle a wide variety of samples^{3,4}, including glycoproteins, integral membrane proteins, proteins and peptides purified by one- or two-dimensional polyacrylamide gel electrophoresis (PAGE) and isoelectric focusing, short hydrophobic peptides, large proteins (molecular weight >100,000), and peptides containing numerous prolyl residues.

The practical success of microsequencing crucially depends on the ability to purify and manipulate small quantities of peptide without either losing the sample or introducing contamination. The failure of this procedure can often be traced to loss of sample during its final purification (perhaps because it adheres to the walls of test tubes), to chemical blockage of the amino-terminus of the peptide by exposure to amino-reactive compounds at the same stage (for example, by reaction with cyanate present in urea solutions), or to contamination by reagents (or their impurities) which may interfere with the Edman chemistry or HPLC analysis of its product (for example, nonvolatile buffer salts or gel components).

Extremely pure reagents and solvents must therefore be used during the preparation and analysis of proteins. HPLC and PAGE (for large peptides and proteins) are suitable. HPLC is the faster and simpler method and causes little damage to the samples, while the development of volatile elution systems (aqueous trifluoroacetic acid solution⁶ and an organic solvent such as acetonitrile or propanol) makes handling of peptides particularly easy because the eluted samples can be applied directly to the sequenator. Many large proteins can now be efficiently chromatographed on the newer 300 µm pore HPLC supports⁷ using these volatile solvents. We have found, using volatile solvents and short (C₃ and C₄) alkylsilane-bonded large pore size silicas, that peptides and proteins in a size range from five amino acids up to >100,000 daltons can be separated in a single run, and that more than 90% of the sample is recovered⁸. In sample purification with PAGE, damage, loss or contamination of the sample are less easily avoided, but electroelution of

Fig. 3 Chemical synthesis of oligonucleotides. The steps involved in the phosphoramidite method are shown: the dimethoxytrityl (DMT) 5' OH protecting group is removed from the silica-bound 3' nucleoside. The nucleoside phosphoramidite is activated with tetrazole and immediately reacted with the growing chain. The resulting phosphite triester is oxidized to the phosphotriester with iodine. This cycle is repeated until the desired chain has been assembled. Protecting groups are removed and the chain cleaved from the silica support, and the target oligonucleotide-purified. Enhanced stability of the nucleoside phosphoramidites makes automated synthesis practical. Yields are typically 95–98% per cycle.



proteins from gels after electrophoresis and subsequent staining with Coomassie blue⁹, or continuous elution of samples from the end of small SDS-PAGE disk gels during electrophoresis (M.H., E. Lujan and L.H., unpublished results), can yield submicrogram amounts of protein. Both methods allow the separation of proteins of very similar molecular weights (<2% difference) and, when necessary, proteins separated by PAGE can be subjected to further purification by HPLC based on their hydrophobic properties.

In principle, it appears that the gas-phase sequencer can analyse samples at the subpicomole range although present limitations on HPLC sensitivity prevent this. One of us (S.K.) has recently introduced¹⁰ a novel class of Edman reagents (Fig. 1) to enhance the sensitivity of the HPLC. These isothiocyanates contain a cryptic chemical function revealed only during the Edman chemistry, which can be exploited to form fluorescent products or other derivatives with enhanced UV or visible absorbance. For example, the sequencing efficiency of 4-(*N*-*tert*-butyloxycarbonylaminoethyl)phenyl isothiocyanate has been shown to be similar to that of phenyl isothiocyanate. The aminomethylphenylthiohydantoins produced with this reagent (after separation by HPLC) can be detected by conventional UV absorbance as well as by the fluorescence of derivatives. Preliminary studies show a severalfold increase of sensitivity compared with conventional phenylthiohydantoins, with detectability well into the subpicomole range (R. Aebersold, L.H. and S.K., unpublished results)¹⁰. We are now working to increase the sensitivity by adapting this reagent for use in our automated sequencers.

Mass spectrometry

In another effort to exploit the potential sensitivity of the gas-phase sequencer, we and engineers at the Jet Propulsion Laboratory (JPL) are developing a sensitive mass spectrometer^{11–13} that will replace our HPLC unit for PTH amino acid analysis. This instrument is small, stable, and simple to operate (Fig. 2). An HPLC-type autosampler is used to deliver aliquots of the sample solution to the tip of a silanized, fused-silica injection probe, where they are dried by evaporation and automatically inserted into the ion source of the mass spectrometer. Ions derived from the sample are accelerated, separated according to mass/charge ratio by the fixed electric and magnetic fields of the mass spectrometer, and focused onto an integrating detector which uses electro-optical methods to generate and amplify a computer-compatible signal representative of the total charge accumulated in each ion beam. The signals generated by several thousand discrete detector elements placed along the ion focal plane are fed into a real-time array processor. The partially processed data are stored on a 40-megabyte hard disk for later access by a minicomputer that calculates the PTH amino acid distribution in each sample before the next one is injected.

The unique aspect of the JPL-Caltech mass spectrometer is that it can be used to monitor simultaneously all ion beams over a wide mass range with sufficient signal-to-noise ratio to detect the arrival of single ions at the focal plane. Its electro-optical ion detection array (EOID, Fig. 2) has 25,000 electron-multiplier detector units per centimetre along the focal plane. Ions impinging on one of the detector elements produce an electron pulse which is then converted to a photon pulse from a phosphor screen (10^6 total amplification). By means of fibre optics, an image of the focal plane is obtained on an array of 5,120 photodiodes, corresponding to a mass range of 25–500 atomic mass units. As the accumulated charge across a capacitor in each photodiode is a measure of the number of ions reaching the corresponding part of the focal plane, the entire mass spectrum can be integrated over times as short as tens of milliseconds up to several seconds.

The sensitivity limits range from 0.5 fmol (PTH-proline) to 10 fmol (PTH-arginine), making the mass spectrometer 100–1,000 times more sensitive than our present HPLC analysis system, while the sample analysis time is <5 min (compared with 30 min), so that one mass spectrometer should be able to handle PTH amino acid samples from several gas-phase sequencers, each with a sample generation time of 50 min. When the EOID mass spectrometer is fully operational, we expect to be able to analyse subpicomole quantities of proteins with the gas-phase sequencers. Then, we should be able to use the analytical technique with highest resolution, two-dimensional PAGE, as a preparative method to isolate proteins for sequence analysis.

DNA synthesis

Rapid, solid-phase methods for synthesizing oligonucleotides have largely replaced the slower, liquid-phase methods used in the pioneering work on chemical DNA synthesis¹⁴. We have incorporated a solid-phase technique using phosphite coupling chemistry^{15,16}, previously developed by one of us (M.C.), into an automated DNA synthesizer¹⁷ using a mechanical design similar to that used in our gas-phase sequencer. The oligonucleotides are synthesized by the stepwise addition of monodeoxynucleoside 3'-phosphoramidites to the 5'-hydroxyl group of a DNA chain, itself covalently linked through its 3'-hydroxyl group to a derivatized, macroporous silica support. The individual steps for each addition are shown in Fig. 3. Once the desired number of monomers has been added to the growing polymer, the completed oligonucleotide is removed from the silica support and the remaining blocking groups are removed from the DNA by treatment with thiophenol and ammonium hydroxide. The correct oligonucleotide is separated by PAGE or HPLC from shorter fragments arising from incomplete couplings.

This automated synthesizer has several interesting features: (1) As oligonucleotides are synthesized on an insoluble polymer

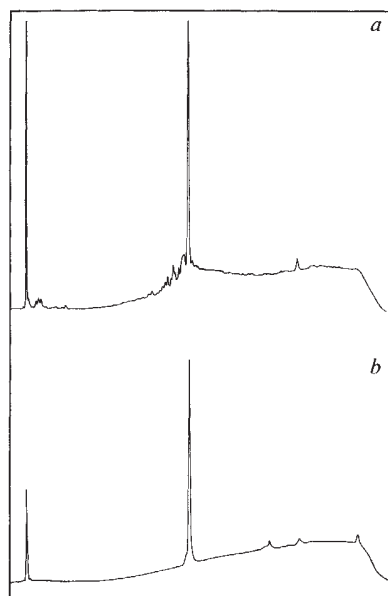


Fig. 4 Reverse-phase HPLC of the product of chemical synthesis of the 36-amino acid residue egg-laying hormone from *Aplysia californica*. The chromatogram in *a* shows the crude product directly after cleavage and deprotection. The purified peptide is shown in *b*. Conditions for both runs: Vydac C₄ 300-Å pore column; flow rate 0.70 ml min⁻¹; buffer A, 0.16% trifluoroacetic acid (TFA) in H₂O; buffer B, 0.10% TFA in 1-propanol; gradient 15–35% buffer B in 60 min; detection at 215 nm. The peptide was also characterized by quantitative sequencing, amino acid analysis, etc.

support, purification is simplified at intermediate steps; excess reagents and most unwanted reaction by-products can be washed from the support by solvents without removing the growing polymer from the reaction vessel. (2) Nucleoside 3'-phosphoramidites are stable compounds which can be stored in solution under dry argon for several weeks without significant deterioration. They are activated immediately before condensation by mixing with a solution of tetrazole in a pre-reaction vessel before being pumped through the main reaction vessel. This *in situ* activation avoids the problem of reagent instability inherent in the use of nucleoside 3'-phosphochloridites as in other automated DNA synthesizers. 'Mixed probes', mixtures of oligonucleotides having identical base sequences at most positions but with two or more bases at others, can be synthesized from appropriate mixtures of phosphoramidites. (3) The process is rapid, each monomer addition cycle lasting only 10 min. (4) Because condensation yields exceed 96% with each of the activated nucleosides, fully automated synthesis of DNA >50 nucleotides long is possible. (5) Reagents and solvents are transferred by pneumatic pressure (argon), not mechanical pumps. This mechanism avoids the blockage by air bubbles that plagues mechanical pumps, causes less physical destruction of the silica support by pressure pulsations and, together with the miniaturized zero-dead-volume valves used in the synthesizer, uses only small quantities of reagents and solvents. (6) The synthesis can be adapted readily to produce quantities of oligonucleotides ranging from 1 to 100 µmol by using reaction vessels of appropriate size.

Peptide synthesis

Peptides can be synthesized chemically using either solution methods¹⁸ or the solid-phase method¹⁹. In solution synthesis, peptides are constructed by successive couplings of short fragments, 3–5 amino acid residues long. At each stage of the synthesis—that is, after each coupling—the intermediate peptides are purified to remove any unreacted precursor peptides before proceeding to the next step. In solid-phase methods¹⁹,

single amino acid residues are added successively to the growing peptide chain which is covalently attached to a solid support by the carboxy-terminal amino acid. Excess reagents can be removed from the growing peptide by simply flushing with solvent, thereby eliminating the need for time-consuming purification of peptide intermediates. However, if the coupling efficiencies at each cycle are not close to 100%, that is, if the coupling does not approach completion, the accumulation of failure sequences severely limits the length of peptide that can be synthesized. In the past, a series of chemical reactions have kept yields at each cycle to ~96%.

Accordingly, we are developing an automated peptide synthesizer based on an optimized solid-phase synthetic chemistry already developed by one of us (S.K., unpublished)^{20,21}. Our process uses a chemically clean, stable derivatized polystyrene/divinylbenzene co-polymer resin for covalent anchoring of the peptide, and uses activated *tert*-butoxycarbonylamino acid derivatives which couple to form the peptide. Coupling efficiencies are greater than 99.8%, a high enough yield to allow the synthesis of long peptides (>70 amino acid residues).

Much of this has been accomplished by recent appreciation of the role of the cross-linked polymer support in keeping protected peptide intermediates in solution and available for reaction^{21–23}. The poor solubility of synthetic intermediates, and the consequent incomplete reactions, is, by contrast, the main problem besetting conventional solution peptide synthesis. This insight has enabled the development of an optimized solid-phase peptide synthesis based on new chemistry, new resin supports and quantitative monitoring of each synthetic cycle. Entire classes of reaction by-products, previously thought to be inherent to the solid-phase approach, have been eliminated, so that there has followed a spectacular improvement in the synthesis of long peptides in high yields and of high product purity (Fig. 4). In most cases, the need for purification of the crude product peptides before preliminary screening for biological activity has been eliminated. Subsequently, the crude peptides can readily be purified by high-resolution ion-exchange or reverse-phase chromatography.

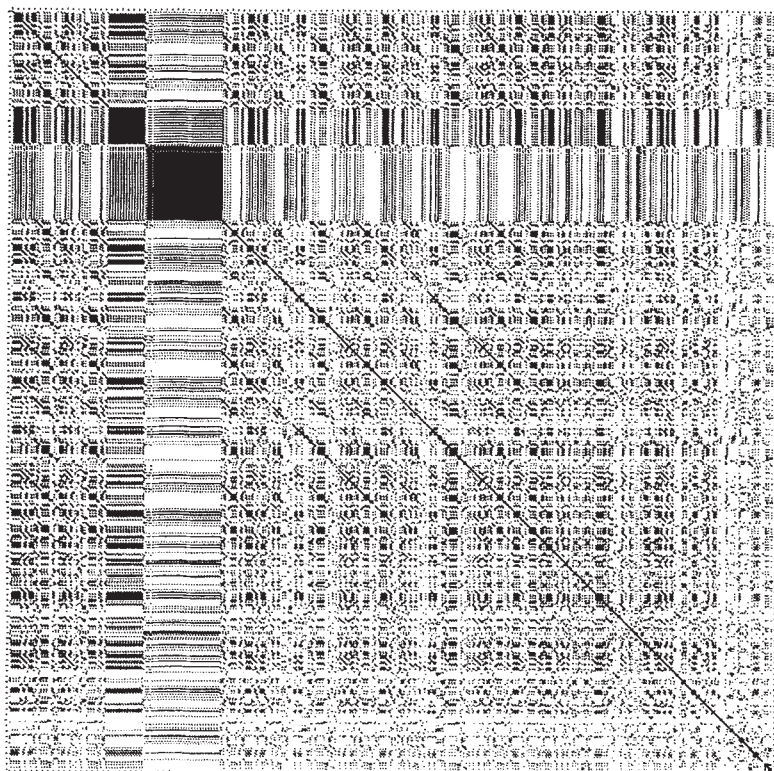
Our automated peptide synthesizer has been designed to exploit both these advances in solid-phase peptide chemistry and the proven mechanical design of our gas-phase protein sequencer and automated DNA synthesizer. Sequence-dependent variations in the synthetic chemistry can be programmed by user-interactive software into each synthesis cycle during the construction of long peptides. The high coupling efficiencies (>99.8% per cycle) and short cycle times (as little as 30 min per amino acid) should allow automated synthesis of 50-residue peptides in 1–2 days.

Computer analysis

As an integral part of our facility, one of us (T.H.) has developed a comprehensive information system to link sources of data (for example, HPLC units for protein sequencing and autoradiograms for DNA sequencing) with various information-processing systems. The core of our system is a Hewlett-Packard 2113E minicomputer with a high-speed processor and a real-time executive operating system (RTE6-V/M) designed to handle large amounts of data flexibly. The system provides automatic control and monitoring of remote instrumentation (such as HPLCs) and permits concurrent access by several users to analytical software as well as various local and national databases by means of a series of linked, interactive, menu-driven control routines. The system contains sufficient hard disk memory so that all the programs and databases are permanently on-line.

We have five linked groups of programs in operation. Spectrophotometric data from HPLCs are digitized, stored and processed by the 3357 Lab Automation System (LAS), a Hewlett-Packard software package. Chromatographic data are further manipulated and converted into a protein sequence by the protein master PROMST, a collection of mathematical and graphic programs designed at Caltech to take the reduced

Fig. 5 A dot matrix analysis of the switch region of the α immunoglobulin gene. A segment of switch region sequence associated with the mouse C_α immunoglobulin gene is compared with itself. A dot has been placed wherever three contiguous bases in one sequence match three in the other sequence. The solid diagonal line represents the line of identity arising from the comparison of the sequence with itself. Any offset diagonals represent direct repeats within the sequence. That this region is constructed primarily of tandemly arrayed 80-base pair homology units is apparent from the series of major offset diagonals spaced at 80-base intervals. Two groups of contiguous exact 5-base pair repeats are responsible for the two box-like features along the identity diagonal.



chromatographic data and turn them into a protein sequence. The DNA master DNAMST is a collection of programs for the input, analysis and display of nucleic acid and protein sequences. DNAMST includes a standard array of routines for sequence file manipulation and primary analyses (translation, restriction enzyme site search, etc.), as well as unique routines for sequence relational studies (alignments, genealogical trees, etc.) and entropy (structural) analysis of particular sequences. The user plot (USRPLT) is a general system for graphic design that can manipulate files from both PROMST and DNAMST to generate technical illustrations. Citation master (CITMST) is a database and retrieval system which can generate reference lists according to specific journal formats for use in laboratory publications.

The importance of computer-aided data analysis to the operation of the Microchemical Facility cannot be overemphasized. Our experience with a pair of two-dimensional matrix analysis programs designed to identify the homology relationships between DNA (or protein) sequences, illustrates this role (Figs 5, 6). One of these (dot matrix analysis²⁴) places the two sequences to be compared on the *x* and *y* axes of a two-dimensional matrix and compares a short nucleotide sequence from the *y* axis with every same-length sequence from the *x* axis (Fig. 5). All matrix positions corresponding to sequence identities are represented by a dot; homologous sequences generate distinct diagonal lines superimposed on a background of randomly placed dots. Two sequences which are homologous except for a gap in one of the sequences generate diagonals which are offset from one another by precisely the number of missing nucleotides. Internal homology relationships are evident as diagonals offset from the major diagonal and can readily be demonstrated when a sequence is compared with itself (see Fig. 5).

Another two-dimensional matrix analysis program, the best-fit matrix routine (T.H., in preparation), is more useful for comparing more distantly related sequences (Fig. 6). The two sequences are again figuratively placed on the *x* and *y* axes of the matrix and a short sequence of nucleotides from the *y* axis compared with every same-length sequence from the *x* axis, but a dot is assigned only to the matrix position corresponding to the high-

est-scored pairing. A match between the central base pair of two aligned segments adds one to the assigned score but other base matches add to the score a value that decreases from one through an arithmetic scale according to the pair's distance from the centre of the segment. Comparison lengths are typically between 11 and 61 bases. Not only can the routine reveal homologies too weak to be detected by the dot matrix routine, but it also filters out most of the distracting background.

Applications

General strategy. Three biotechnologies—recombinant DNA, monoclonal antibodies and microchemical instrumentation—which have been developed in the last 10 years, have had, and will continue to have, a profound impact on our ability to tackle fundamental biological problems. As illustrated in Fig. 7, these techniques are synergistic in that the scientist can start at any point (protein, messenger RNA (mRNA) or genes) and use the biotechnologies in combination to generate genes from proteins or proteins from genes and at the same time to produce highly specific reagents to identify gene products. Thus, there is almost complete symmetry in our ability to approach the isolation of genes and the identification of proteins at any point on this circle. **Rare-message genes.** The unique capabilities of the Microchemical Facility are particularly suited to the study of genes from which only small quantities of mRNA are transcribed, including those coding for most receptor molecules, growth and differentiation factors, and hormones. These are difficult to isolate by conventional techniques that use biological assays to screen expression products derived from the mRNAs since they may only be present at levels of 0.1–0.0001%; they may, however, be identified with the Microchemical Facility (see Fig. 8) as described below.

Microgram quantities of purified proteins or peptides are subjected to amino acid sequence analysis using the gas-phase microsequencer and 5–7-residue stretches of the amino acid sequence are translated by the genetic code to identify all possible originating DNA sequences. One to three stretches with minimal ambiguity at the DNA level are selected, and all of the 14–20-mer oligonucleotides encoding these regions are made

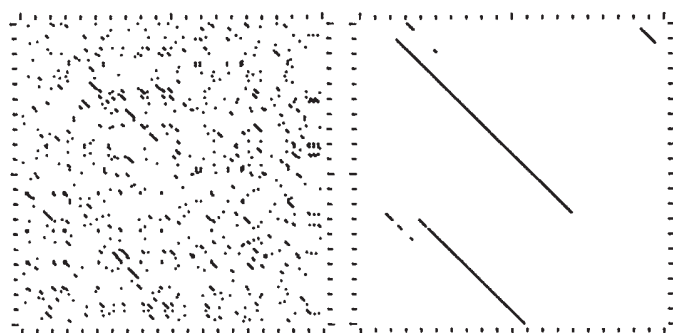


Fig. 6 Comparisons of exon 4 of the mouse C_μ immunoglobulin gene (y axis) with exon 4 of the $H-2L^d$ class I gene (x axis) using the dot matrix (see Fig. 5) and best-fit matrix routines. The dot matrix (left) marks with a dot each identity of three contiguous bases. The best-fit comparison (right) plots the best score for pairings of 51 bases (25 bases on either side of aligned base pairs). The two major lines not only indicate homology between the two genes, but also homology between the first and second halves of the exons. Clearly, the best-fit matrix analysis eliminates much of the noise present in the dot matrix analysis in comparing these two distantly related gene fragments (see text).

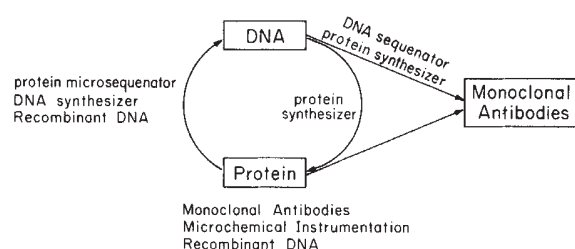


Fig. 7 A diagrammatic representation of the synergy of three biotechnologies—recombinant DNA technology, microchemical instrumentation and monoclonal antibodies.

using the automated DNA synthesizer. These oligonucleotides are used to screen complementary DNA (cDNA) libraries containing the desired gene^{25,26}. Alternatively, the 3'-most oligonucleotide mixture may be used as a specific primer to generate a cDNA library which can then be screened with the remaining synthetic probes.

If the protein amino acid sequence data available from the microsequencing are sufficient to provide several reasonably unambiguous DNA probe mixtures, it may be possible to clone rare-message genes directly from genomic rather than cDNA libraries, despite the more than 100-fold increase in complexity of the typical genomic library, by successive screening of clones positive to one probe, with other probes. To screen oligo(dT)-primed cDNA libraries for rather large gene transcripts, or mRNAs containing large untranslated regions at the 3' end, as well as for specific priming of reverse transcription, it is important that oligonucleotide probes should be located as close as possible to the 3' end, corresponding to the C-terminal region of a long protein. We are improving our chemical and enzymatic techniques for fragmenting small quantities of proteins so that fragments near the C-terminus can be microsequenced. This should also yield more extended sequence data, essential to finding amino acid sequence regions defining oligonucleotide probes of minimal ambiguity, and will be useful for defining the start, end and intron/exon boundaries of genes for which we do not have mRNAs.

Gene synthesis. As the DNA synthesizer can be used to make oligonucleotides up to 50 bases long in one operation, it can be used to synthesize complete genes relatively quickly. This is done by synthesizing overlapping fragments from the two

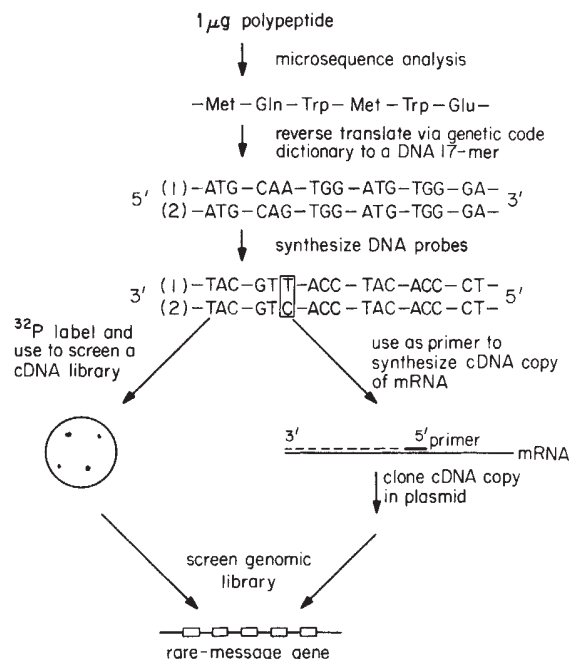


Fig. 8 Strategy for the cloning of rare-message genes. The protein sequence is determined and reverse-translated into mRNA and DNA. A region of minimal ambiguity, that is, consecutive codons having as little degeneracy as possible, is identified and a corresponding collection of heterogeneous DNA fragments of 14–20 nucleotides are synthesized. These synthetic oligonucleotide probes can be used to screen directly a cDNA (or genomic) library or they can be used as primers to initiate specific cDNA synthesis.

strands of the double-stranded gene in such a way that its fragments can hybridize with one another and can be ligated together correctly. Using this approach, one of us (S.H.) has synthesized a synthetic gene coding for *Escheria coli* serine tRNA in less than 2 weeks. This capability may be important for the production of mutant or hybrid genes and of genes in which alternative codons have been chosen to make the best use of transfer RNA levels in particular bacteria, yeast or mammalian cells.

Mutant genes also can be generated from parent genes using short strands of synthetic DNA in a process known as oligonucleotide-directed mutagenesis²⁷. This procedure allows the modification, at specified amino acid residues, of genes encoding enzymes, hormones or other proteins, to produce mutant proteins rapidly and to order. The synthetic oligonucleotides are used not only to define and create the mutation, but also to select the desired mutant bacterial plasmids. Use of the techniques permits the type of systematic variation of protein structure that can allow precise elucidation of the structure-function relationships governing the action of a protein. Mutant genes made by this procedure or by total gene synthesis can then be combined with appropriate transcriptional control elements and inserted into microorganisms for large-scale production of the new proteins.

Peptide antibodies. Although Southern blot analysis can identify multigene families and DNA-mediated gene transfer experiments can be used to determine which of them express products, often there are no reagents that will identify the translation products of the genes. In these circumstances, individual DNA sequences can provide the corresponding protein sequences that define hydrophilic peptide regions unique to each of the products; these are synthesized, conjugated to carrier proteins, and used to immunize animals to give conventional heterogeneous antisera^{28,29} or monoclonal antibodies³⁰. Many of these antibodies will recognize the parent protein, in the correct experimental conditions, and so will have predetermined specificity

for certain sites that uniquely define the different members of the multigene family. This approach is being followed with the class I genes of the major histocompatibility complex in the mouse, where a single class I cDNA probe cross-hybridizes with at least 32 distinct genes^{24,31}. The antibodies produced by this technique can be used to identify the cells in which the proteins are expressed and to assist in the determination of their function. Such reagents are also useful in determining post-translational processing of proteins²⁹.

The peptide antibody technique is also potentially valuable in determining whether an amino acid sequence determined by microsequencing derives from a protein possessing a specific biological activity or from a co-purified contaminant; this is particularly important in the case of hormones or growth factors when purification requires enrichment of a million times or more. Once part of the amino acid sequence of an apparently pure protein has been determined, it can be used to synthesize peptides for producing antibodies, which can then be tested for their ability to remove the appropriate biological activity from a protein solution containing it.

Peptide domains. It should soon be possible to synthesize peptide fragments up to 100 amino acids long in less than a week using the synthetic methodology described previously, either alone or in conjunction with the peptide splicing chemistry now being developed. If so, it should be possible to synthesize complete peptide domains (native or analogues) and to test their structure-function properties. This procedure could well be more rapid and versatile than the genetic engineering of the corresponding genes, followed by transcription and translation.

Microsequence screening. In the future, it will be possible to isolate from two-dimensional PAGE³² or comparable protein separation procedures, a large series of proteins that can be screened by amino-terminal microsequencing to identify individual proteins. This rapid structural screening could be used to locate a specific protein or to define the interrelatedness of a series of proteins.

Enhanced DNA sequencing. The ability to synthesize rapidly short fragments of DNA provides a means of improving the efficiency and speed of DNA sequencing³³. At present, a short strand of DNA, complementary to part of a cloning vector such as M13 (ref. 34), is used to prime second strand synthesis in the dideoxy sequencing method³⁵. Because of technical limits associated with various stages of the method, one can generally obtain the sequence for only 200–300 nucleotides for a given piece of DNA before some of the base assignments become ambiguous. In order to define the sequence of a large piece of DNA, one has to generate in cloned form a large number of relatively small fragments, sequence as much of each as possible, and then order the fragments into the full-length stretch of DNA. We have carried out sequencing using a 3-kilobase (kb) length of DNA cloned in M13 (or a similar vector designed to handle inserts larger than the 2–3 kb that work well with M13), and

sequenced as far as possible using the M13 primer. Then a new primer, complementary to a region near the end of the newly defined sequence, was synthesized to begin a second round of sequencing by priming at this point (J. Kobori, E. Strauss and L.H., unpublished observations). By repeating the sequencing–new primer synthesis approach, one can sequence large stretches of DNA without the time-consuming process of subcloning and ordering small fragments. Moreover, the cDNA strand can be sequenced with great speed by synthesizing an appropriate number of complementary synthetic oligonucleotides.

Synthetic DNA as diagnostic probes. The identity of bacteria or viruses responsible for infections is often confirmed by immunoassay procedures that recognize specific antigenic determinants on proteins produced by the agents and present in the host. Genetic diseases arising from inheritable alterations in the genes encoding vital proteins are also characterized by analysis of the protein products. In both cases, however, a direct examination of the genetic information associated with the disease can often be done at an earlier stage in the disease process than is possible using classical immunoassay procedures. The ability to synthesize short strands of DNA, complementary to the DNA from the microorganisms or the mutant host genes, should allow diagnostic applications that analyse DNA rather than proteins to become generally used methods. Diagnosis of infections of the central nervous system and prenatal screening of inherited disorders are well suited to this approach. The feasibility of the use of genetic probes has already been demonstrated in the case of inherited haemoglobinopathies³⁶.

Conclusion

The Microchemical Facility provides enormously sensitive and powerful approaches to tackling a variety of difficult problems in molecular and developmental biology, and in neurobiology. We have discussed the characteristics of the instruments and pointed out various synergistic strategies by which they can be used in combination with recombinant DNA and hybridoma techniques. We are presently developing an automated DNA sequencer that will complete a series of instruments designed for synthesis and analysis of both genes and proteins. The speed and simplicity of these techniques, used together with recombinant DNA and monoclonal antibody technology, will provide unparalleled opportunities to resolve fundamental problems in biology.

We acknowledge the support of the Weingart Foundation for the development of the Microchemical Facility, which is housed in the Braun Laboratories for Chemical Biology, and The Monsanto Corporation and NIH for supporting the development of the mass spectrometer. A portion of the funding for continued operation of the Microchemical Facility is provided by the NCI. P.T. is a senior research assistant of the Belgian National Fund for Scientific Research.

- Edman, P. *Acta chem. scand.* **10**, 761–768 (1956).
- Edman, P. & Begg, G. *Eur. J. Biochem.* **1**, 80–91 (1967).
- Hewick, R. M., Hunkapiller, M. W., Hood, L. E. & Dreyer, W. J. *J. Biol. Chem.* **256**, 7990–7997 (1981).
- Hunkapiller, M. W. & Hood, L. E. *Science* **219**, 650–659 (1983).
- Tarr, G. E., Beecker, J. F., Bell, M. & McKean, D. J. *Analyt. Biochem.* **84**, 622–627 (1978).
- Bennett, J. P. J., Hudson, A. M., McMartin, C. & Purdon, G. E. *Biochem. J.* **168**, 9–13 (1977).
- Lewis, R. V., Fallon, A., Stein, S., Gibson, K. D. & Udenfriend, S. *Analyt. Biochem.* **104**, 153–159 (1980).
- Tempst, P., Hunkapiller, M. W. & Hood, L. E. *Analyt. Biochem.* **137**, 188–195 (1984).
- Hunkapiller, M. W., Lujan, E., Ostrander, F. & Hood, L. E. *Meth. Enzym.* **91**, 227–236 (1983).
- L'Italien, J. J. & Kent, S. B. H. 2nd int. Symp. on HPLC of Proteins, Peptides, and Polynucleotides, Baltimore, Maryland, December 1982, Abstr. 600; *J. Chromatogr.* **283**, 149–156 (1984).
- Giffin, C. E., Boettger, H. G. & Norris, D. D. *Int. J. Mass Spectrom. Ion Phys.* **15**, 437–449 (1974).
- Dreyer, W. J. *et al. Clin. Chem.* **20**, 998–1002 (1974).
- Boettger, H. G., Giffin, C. E. & Norris, D. D. in *Monograph on Image Detectors in Chemistry* (ed. Talmi, Y.) 291–318 (American Chemical Society, 1979).
- Khorana, H. G. *et al. J. molec. Biol.* **72**, 209–217 (1972).
- Matteucci, M. D. & Caruthers, M. H. *J. Am. chem. Soc.* **103**, 3185–3191 (1981).
- Beaucage, S. L. & Caruthers, M. H. *Tetrahedron Lett.* **22**, 1859–1862 (1981).
- Horvath, S. J., Firca, J., Hood, L., Hunkapiller, T. & Hunkapiller, M. W. *J. Biol. Chem.* (submitted).
- Yajima, H. & Fujii, N. in *Chemical Synthesis and Sequencing of Peptides and Proteins* (eds Liu, T. Y., Schechter, A. M., Heinrikson, R. L. & Condliffe, P. G.) 21–39 (Elsevier, Amsterdam, 1981).
- Barany, G. & Merrifield, R. B. in *The Peptides* Vol. 2 (eds Gross, E. & Meienhofer, J.) 1–284 (Academic, New York, 1980).
- Kent, S. B. H., Mitchell, A. R., Engelhard, M. & Merrifield, R. B. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2180–2184 (1979).
- Kent, S. B. H. in *Biomedical Polymers* (eds Goldberg, E. P. & Nakajima, A.) 213–242 (Academic, New York, 1980).
- Sarin, V. K., Kent, S. B. H. & Merrifield, R. B. *J. Am. chem. Soc.* **102**, 5463–5470 (1980).
- Kent, S. B. H. & Merrifield, R. B. in *Peptides 1980, Proc. 16th European Peptide Symp.* (ed. Brunfeldt, K.) 328–333 (Scriptor, Copenhagen, 1981).
- Steinmetz, M. *et al. Cell* **24**, 125–134 (1981).
- Noyes, B., Mevareck, M., Stein, R. & Agarwal, K. L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1770–1774 (1979).
- Steinmetz, M. *et al. Nature* **300**, 35–42 (1982).
- Dalbadie-McFarland, G. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 6409–6413 (1982).
- Walter, G., Scheidtmann, K.-H., Carbone, A., Laudano, P. & Doolittle, R. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5197–5200 (1980).
- Sutcliffe, J. G. *et al. Nature* **287**, 801–805 (1980).
- Lerner, R. A. *Nature* **299**, 592–596 (1982).
- Steinmetz, M., Winoto, A., Minard, K. & Hood, L. *Cell* **28**, 489–498 (1982).
- O'Farrell, P. H. *J. Biol. Chem.* **250**, 4007–4021 (1975).
- Hamlyn, P. H., Gait, M. J. & Milstein, C. *Nucleic Acids Res.* **9**, 4485–4494 (1981).
- Messing, J., Crea, R. & Seeburg, P. H. *Nucleic Acids Res.* **9**, 309–321 (1981).
- Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. J. *J. molec. Biol.* **143**, 161–178 (1980).
- Suggs, S. V., Wallace, R. B., Hirose, T., Kawashima, E. H. & Itakura, K. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6613–6617 (1981).

Multiple imaging of quasars by galaxies and clusters

Ramesh Narayan^{*†}, Roger Blandford[†] & Rajaram Nityananda[‡]

[†] Theoretical Astrophysics, California Institute of Technology, Pasadena, California 91125, USA

[‡] Princeton University Observatory, Princeton, New Jersey 08544, USA and Raman Research Institute, Bangalore 560080, India

Multiply-imaged quasars have generally been explained as a combination of gravitational lensing by galaxies and magnification by clusters. The observational difficulties with this view suggest that some such quasars may be more naturally interpreted as images split by the cores of very rich clusters. The observational distinctions between the two models are discussed.

GALAXIES at large redshift can behave as gravitational lenses and produce multiple images of background quasars^{1,2}. Five convincing examples of lensing action have recently been discovered^{3-7,35} (Q0957+561, Q1115+080, Q2345+007, Q2016+112, Q1635+267). However, the observed image separations, which range up to 7 arcs, are conspicuously larger than the 1–2 arc separations expected from normal galaxies. Thus, it has generally been concluded that galaxy lensing action is considerably assisted by the magnification of an associated cluster⁸. The cluster also amplifies the quasar flux and so the distribution of observed image separations is biased towards large values (ref. 9, and J. P. Ostriker, personal communication).

However, problems remain. In two cases, no lensing galaxies have been found near the images. Furthermore, only in Q0957+561 do we see the odd number of images that we would expect from a transparent lens¹⁰, and even in this case the identification is ambiguous¹¹. Also, in two of the three cases where we do see galaxies, the image geometries are hard to model. These difficulties motivated us to examine an alternative view in which the lensing action is primarily attributable to the cluster (as suggested in ref. 12 for a particular case), and the galaxy, if present, is incidental. This explanation is more than just a change of emphasis because it leads to different predictions regarding image orientations and parities as well as the location of the missing images.

Cluster-assisted imaging by galaxies

The standard, simple image geometry is shown in Fig. 1a. Bright images are found on either side of a giant galaxy and a faint third image (which is missing in most cases) is presumed to be located near the galaxy nucleus. The image separations and fluxes are magnified by the gravitational field of a rich cluster surrounding the galaxy. For the moment, we assume circular symmetry and ignore distortion and the possible creation of additional images¹³. We follow ref. 9 and model the galaxy as an isothermal sphere with one-dimensional velocity dispersion ν and the cluster as a sheet of uniform surface density Σ . Rays with an impact parameter exceeding the galaxy core radius but less than the outer radius are bent through an angle $\alpha = 4\pi\nu^2$ ($c = G = H_0 = 1$ henceforth; we use Hubble's constant, $H_0 = 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$.) If the quasar and the lensing galaxy are at redshifts z_q and z_l , then the observed angular separation of the two images is

$$\theta = \frac{2\alpha d(z_l, z_q)}{d(0, z_q) - 4\pi\Sigma d(0, z_l) d(z_l, z_q)} \quad (1)$$

with $d(z', z)$ denoting the angular diameter distance of redshift

z from z' . (In this model the cluster core, also at z_l , simply provides an isotropic linear magnification of $[1 - 4\pi\Sigma d(0, z_l)d(z_l, z_q)/d(0, z_q)]$, assumed to be positive for the moment.) Given a quasar redshift z_q , an image separation θ , and a cosmological model, we can calculate the required $\nu(\Sigma, z_l)$ of the lensing galaxy. We then relate ν to the galaxy luminosity L using the Faber–Jackson¹⁴ relation, $\nu = 1.5^{1/2}\nu^*(L/L^*)^{1/4}$ where $\nu^* (= 250 \text{ km s}^{-1}$, ref. 15) and L^* (blue magnitude $M_B^* = -19.1$, ref. 16) correspond to a fiducial galaxy, and we have included a factor $1.5^{1/2}$ to allow for the underlying halo having a higher velocity dispersion than the galaxy. (It is not clear that galaxies in cluster cores do have haloes and, even if they do, the factor 1.5 is probably an overestimate for the brightest cD galaxies; however, reducing this factor only strengthens our conclusion.) The luminosity L is converted to a r magnitude using standard colours and k -corrections¹⁷.

Let us fix the cluster density Σ and consider a given image separation θ and quasar redshift z_q . If the lensing galaxy is at small z_l the bending angle is effectively constant at $\theta/2$ and the flux received from the lens $\rightarrow \infty$ as $z_l \rightarrow 0$. On the other hand, as $z_l \rightarrow z_q$, the bending angle increases rapidly and the larger ν that this implies again leads to a diverging flux. There is thus an intermediate range of redshift, $z_{\min} - z_{\max}$, within which the optical magnitude of the lens will exceed a given value m , and in particular a maximum lens magnitude. Next we adopt differential luminosity functions for the galaxies¹⁸ and the background quasars¹⁹, $d\eta_g/dL \propto L^{-5/4} \exp(-L/L^*)$ and $d\eta_q/dL \propto L^{-3}$. We can now compute the probability $P_m(>m; \Sigma, \theta, z_q)$ that the r magnitude of the lensing galaxy exceeds the observed or upper limit value m for a given value of the surface density Σ .

$$P_m(>m; \Sigma, \theta, z_q)$$

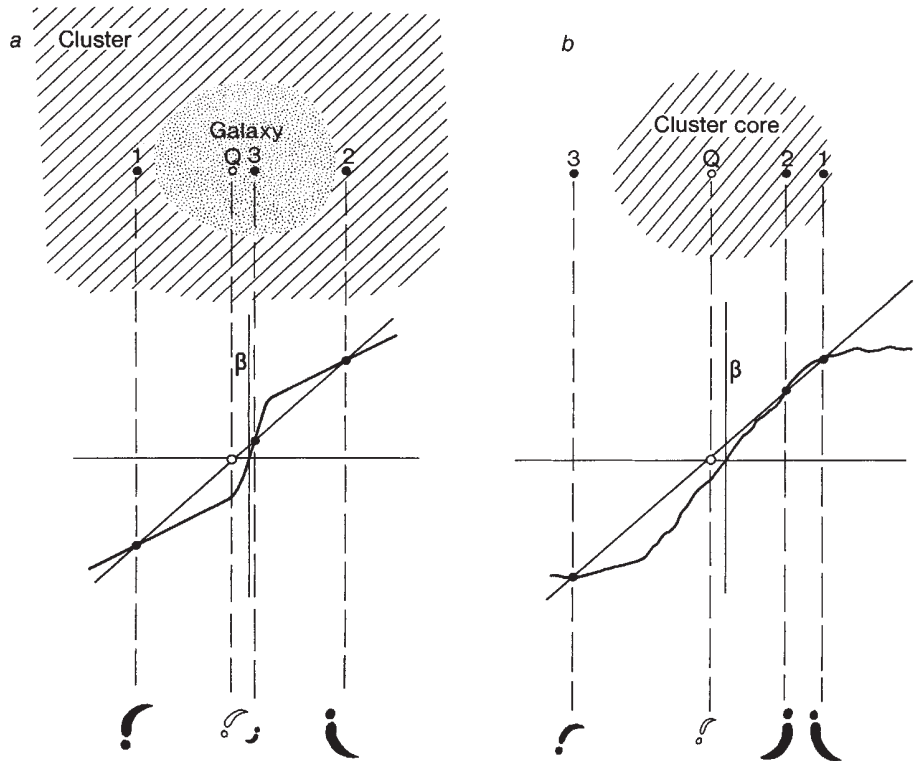
$$\propto \int_{z_{\min}}^{z_{\max}} \left(\frac{L^*}{L}\right)^{5/4} \exp\left(-\frac{L}{L^*}\right) \left(\frac{dL}{d\theta}\right) A^4 \left[\frac{\theta^2 d^2(0, z_l)}{A^2}\right] \times (1+z_l)^3 \left(\frac{dt}{dz_l}\right) dz_l \quad (2)$$

Here A is the linear amplification due to the cluster and t is cosmic time. The factor A^4 arises from the quasar luminosity function ($\eta_q(>L) \propto L^{-2}$), noting that flux is amplified by A^2 , and the A^2 in the denominator allows for the reduced geometrical cross-section (for a given θ) in the presence of cluster magnification⁹.

We have calculated $P_m(>m; \Sigma, \theta, z_q)$ for the five known examples of lensing (Table 1). The conclusion is quite clear. Except in the case of the double quasar Q0957+561 (for which the observed lens redshift is close to where it is most likely to be, given the quasar redshift and the image separation, see also

* Permanent address: Raman Research Institute, Bangalore 560080, India.

Fig. 1 Schematic illustration of the two contrasting views of gravitational lensing. *a*, Imaging by a giant galaxy assisted by the magnification of a surrounding cluster. The two bright images 1 and 2 should be found on either side of the galaxy and the faint image 3 should be within the galaxy core. *Q* is the true position of the quasar. Additional pairs of images may be created if the gravitational field is non-circular. The graph shows the reduced deflection, $\beta = (d(z_i, z_q)/d(0, z_q))$ times the actual deflection at the lens, as a function of angular displacement from the galaxy nucleus. Images are located at positions where this intersects a line of unit gradient passing through the undeflected position of the quasar. The predicted shapes of bent core-jet radio sources are illustrated at the bottom. *b*, Imaging by the core of a rich cluster. Note the difference in scale compared with *a*. In the most prominent examples, the two bright images should be found on one side of the core with the faint third image diametrically opposite. Large amplification bias occurs when the bright images are nearly merging as shown in the graph of reduced deflection. (The corrugations, caused by the galaxy-sized granularity of the potential well, limit the decreasing image separation and amplification.) The radio source shapes predicted for the two bright images should be distinguishable from those of case *a*.



ref. 20), lensing action by a sufficiently faint galaxy requires the assistance of a cluster that has a surface density that is a substantial fraction of the critical value Σ_c to produce multiple images by itself. (The focusing power of a given cluster core depends on its redshift; Σ_c refers to the critical density at the optimal redshift, which is typically ~ 0.5 .) As there seems to be no reason why the distribution of cluster surface densities should cut off just below Σ_c , we suggest that some of the observed multiple quasars are produced entirely by clusters or cluster-like potential wells.

The absence of third (or, in the case of Q1115+080, fifth) images is perplexing. In the simple, galaxy-lensing geometry (Fig. 1*a*), the faint image should be located between the observed images with a flux a factor $\sim (\theta_g/\theta)^2$ fainter than the sum of the fluxes from the two brightest images, where θ_g is the angular radius of the core of the galaxy and θ is the separation of the bright images. The published lower limits on the fluxes of the faint images can then be converted into upper limits on the core radii of the galaxies. These are summarized in Table 1. We see that the necessary core radii are just compatible with those measured for brightest cluster galaxies, for example, 1.3 kpc in ref. 21, except for Q1115+080 where there seems to be a discrepancy.

The third difficulty associated with invoking galaxies assisted by clusters to produce multiple images is that it is not always easy to reproduce the observed image configurations. The image positions in Q0957+561 are relatively easy to reproduce by combining a galaxy and a cluster^{8,22}. The four observed images in Q1115+080 are also readily reproduced using an elongated galaxy located within them^{19,23}; however, the only reported galaxy in the vicinity of these images is located outside the triangle formed by images A, B and C²⁴. In the case of Q2016+112, the A and B images and the galaxy form a triangle with angles $\sim 90^\circ, 60^\circ$ and 30° respectively. We have attempted to model this geometry using an elliptical distribution of surface density and find that a 60° angle at the galaxy is invariably

accompanied by similar angles at the images. This seems to rule out a simple galaxy model for this case. No constraints are imposed in the other two known cases of lensing, as galaxies have not been observed.

Galaxy-limited lensing by clusters

We now contrast the creation of multiple images by galaxies with that by clusters. We assume that the surface mass density in the cluster core can be described by a parabolic variation with $\alpha = \alpha_0$, $r > 1.155r_c$. Our conclusions are not very sensitive to $\alpha_0[1.30(r/r_c) - 0.325(r/r_c)^3]$, where the core radius r_c is the point at which the surface density falls to half the central value. We match the core smoothly on to an external 'isothermal sphere' with $\alpha = \alpha_0$, $r > 1.155r_c$. Our conclusions are not very sensitive to the precise density distribution assumed. Clusters with central density Σ greater than $\Sigma_c = d(0, z_q)/4\pi d(z_1^*, z_q)$, where z_1^* is the optimal redshift for cluster lensing, can produce lensing without the aid of an additional galaxy.

It is usually assumed that clusters produce image splittings on the scale of their angular core radius θ_c , that is ~ 30 arc s. Paczynski and Gorski¹² suggested cluster lensing as an explanation for three quasars found within a couple of arc minutes of each other with nearly the same redshift. Two other examples, Q0307-195 (ref. 25) and Q0107-025 (ref. 26), could also have a similar explanation. However, we note that clusters need not always produce large image splittings. When two of the images are on the point of merging, that is, image separation $\theta \ll \theta_c$, the amplification will be much larger (a graphic small scale illustration of this phenomenon is provided by the split A image in Q1115+080 which is three magnitudes brighter than the B and C images), and these events will be strongly favoured by selection. Mergers tend to occur on one side of the cluster, offset from the centre by an angle $\sim \theta_c$, with a fainter third image at $\sim \theta_c$ on the opposite side. Defining $\epsilon = (\Sigma - \Sigma_c)/\Sigma_c$ and letting $\Delta\theta'$ be the offset of the true quasar position from the location

Table 1 Results for the known examples of gravitational lensing

Quasar	θ (arc s)	z_q	m_r	Σ/Σ_c	z_l	Flux ratio	r_c (kpc)
Q0957+561	6.2	1.41	18.5	0.16	0.37 ± 0.07	0.014‡	<1.1
Q1115+080	2.3	1.72	>22*	>0.35	0.62 ± 0.10	<0.018§	<0.8
Q2345+007	7.3	2.15	>23†	>0.87	0.48 ± 0.05	<0.013	<2.0
Q2016+112	3.4	3.27	23.2	0.66	0.73 ± 0.17	—	—
Q1635+267	3.8	1.96	>23.5	>0.81	0.49 ± 0.05	<0.015	<1.1

Σ/Σ_c is the value for which there is 50% probability of having a lens magnitude greater than the observed m_r , in a Friedman universe with $\Omega = 1$. z_l is the distribution of lens redshift for this case. The core radius r_c of the lensing galaxy is estimated using the ratio of the flux in the faint third/fifth image to the flux in the other bright images.

* From ref. 33; † from ref. 34.

‡ Based on VLBI observations¹¹.

§ We have considered only images B and C since the merging pair A_1A_2 may be anomalously bright.

θ' for which the close images merge and disappear, the separation θ and amplification A of the close pair of images are given by

$$\theta = 2.2\varepsilon^{-1/2}(\theta_c\Delta\theta')^{1/2} \quad (3)$$

$$A = 0.81\varepsilon^{-5/4}(\theta_c/\Delta\theta')^{1/2} = 1.7\varepsilon^{-3/2}(\theta_c/\theta) \quad (4)$$

The third image, located approximately a cluster diameter away, has $A \sim \varepsilon^{-3/2}$.

It is apparent that the strongest magnification of the merging images occurs when $\theta \ll \theta_c$ and $\varepsilon \ll 1$. The differential cross-section $d\sigma(\theta)$ for image separations between θ and $\theta + d\theta$ is $d\sigma/d\theta = 2\pi\theta'(d\Delta\theta'/d\theta) = 2.1\varepsilon^2\theta$. If we continue to assume that the differential quasar luminosity function has slope = 3, then the effective cross-section σ' per unlensed quasar brighter than the flux limit, after correcting for amplification bias, is $d\sigma'/d\theta = A^2(d\sigma/d\theta) = 6.3\theta_c^2/(\varepsilon\theta)$. The cumulative effective cross-section for image separations less than a given θ is

$$\sigma' = 6.3(\theta_c^2/\varepsilon) \ln(\theta/\theta_{\min}) \quad (5)$$

where the minimum separation $\theta_{\min}$ is discussed below. Thus we expect roughly equal numbers of cases of multiple imaging with equal logarithmic intervals of image separation, and the small image separations of the known cases ($3 \leq \theta \leq 7$ arc s) are understood as a natural consequence of amplification bias.

We do not know the distribution of ε in the supercritical clusters we are considering. However, the above analysis shows that, as a further consequence of amplification bias, $d\sigma'/d\varepsilon \propto 1/\varepsilon$. The logarithmic divergence again ensures that most multiple images in a magnitude limited sample should be formed by clusters that have surface density barely above critical.

Two effects determine the cutoff $\theta_{\min}$ in the image separation. First, the steep luminosity function must flatten at the faint end. Second, the potential of the cluster is expected to have inhomogeneities on the galaxy scale. These will produce fluctuations $\delta \sim 1$ arc s in the cluster deflection. Such galaxy scale roughness cuts off the distribution of bright merging pairs on angular scales $\leq (\delta\theta_c)^{1/2} \sim 5$ arc s. A similar argument gives the cutoff in ε to be $\sim (\delta/\theta_c)^{2/3} \sim 0.1$. We further note that, while in a completely smooth cluster potential the intensities of a pair of merging images become asymptotically equal, this need not be true once roughness of the cluster potential on a scale smaller than the image separation is taken into account. Indeed, as the most prominent lenses should be the marginal ones, we might expect to find several cases in which a giant galaxy provides the final deflection necessary for multiple images.

For simplicity and clarity, we have discussed the radial merger of two bright images near the boundary of the cluster core. Tangential mergers of bright pairs produced by deviations from circular symmetry are also possible, the classic example being the galaxy lensing models proposed^{11,23} for Q1115+080. In this type of event, one expects three fainter images (instead of one), again separated by roughly a core diameter from the bright pair. The above estimates for amplifications and cross-sections con-

tinue to hold for this case although the numerical factors will be different.

The two merging images in cluster lensing have opposite parity, as do the bright images in galaxy lensing. However, the orientations of position angles in the two types of lensing are distinct (Fig. 1). This is relevant to those compact radio quasars that contain a core-jet structure. Very long baseline interferometry (VLBI) should easily distinguish between the two possibilities. Q0957+561, the only case where VLBI observations have been made, is in better agreement with the galaxy lensing scenario, although the parities of the two images seem to be the same instead of opposite.

How likely are rich clusters to lens quasars? The condition that a cluster at an optimal redshift ~ 0.5 be able to produce multiple images, $\Sigma > \Sigma_c \approx 0.8 \text{ g cm}^{-2}$, translates into the requirement²⁷ that the one-dimensional velocity dispersion exceed $\sim 1,000$ (core radius/100 kpc)^{1/2} km s⁻¹. Typical local rich clusters have velocity dispersions about half this critical value and several, such as A401, A426, A1775, A2029 and A2256, appear to be supercritical^{28,29}. However, the core radii of clusters may not be well measured by the galaxy distribution—for instance, X-ray observations of clusters generally suggest smaller radii³⁰—and many more clusters may be supercritical than hitherto imagined. For illustration, we adopt the local integral distribution function for clusters with surface density in excess of Σ to be $\eta_c(>\Sigma) \propto \Sigma^{-3}$ and assume $\Sigma \propto \nu^4$. Then, using the above estimates for the cross-sections in a $\Omega = 1$ Friedman universe (the results are insensitive to the choice of the exponent and cosmological model), we find that we need ~ 300 supercritical clusters Gpc⁻³ to produce a fraction 0.001 of lensed quasars in a magnitude limited sample. This compares favourably with the estimated local density of clusters richer than Coma ($\sim 3,000 \text{ Gpc}^{-3}$, ref. 31) and indicates that our cluster lensing model could be competitive with the galaxy lensing quantitatively studied in ref. 9.

Observational implications

We have argued that amplification bias is just as likely to be important in lensing by clusters as it is in lensing by galaxies. Consequently, quasars are just as likely to be multiply-imaged by clusters as by galaxies. Moreover, because of amplification bias, image separations of a few arc seconds are highly probable in cluster lensing despite the large angular size of cluster cores. We feel that although Q0957+561 is a convincing example of cluster-assisted galaxy imaging, the other four examples are equally well interpreted as cluster imaging. Fortunately, this idea is readily tested in these four (and indeed any future) cases. First, we should seek the lensing cluster, whose core should lie to one side of the images rather than surround them. This is reportedly true^{6,13} of Q1115+080 and Q2016+112 and distant galaxies have been detected³² in the vicinity of Q2345+007. The fact that these clusters are not as rich as one might expect is, of course, a matter for concern. The most probable redshift for the cluster depends on the quasar redshift but will generally lie in the range 0.3–1. Second, we predict that a third image

should be found on the other side of the cluster, typically two or three magnitudes fainter than the existing bright images. (The search for this may be accomplished most effectively by using a filter centred on one of the bright lines in the bright images.) Third, we can examine the image geometries of compact radio quasars using VLBI and see if the source shapes in the two brightest images are more nearly reflections about a mirror perpendicular to the line joining the images rather than parallel to this line as is the case for galaxy lensing (Fig. 1).

There are other possible types of lenses. Some compact groups of galaxies appear to have surface densities approaching Σ_c and there is no objection of principle to the existence of cluster-like potential wells of dark matter (or equally plausibly strings) not

illuminated by the light of constituent galaxies. In any case, we note that cluster size lenses can act as telescopes on a cosmological scale and enable us to explore the faintest and most distant quasars, galaxies and radio sources. Searches for background objects in the field of rich, centrally-condensed clusters in the redshift range 0.3–1 are of particular interest.

We thank R. Gott, C. Lawrence, J. Ostriker, B. Paczynski and D. Schneider for stimulating discussions of the properties of gravitational lenses. Financial support by NSF under grant AST 82-13001, and the Alfred P. Sloan Foundation is acknowledged. R. Nityananda thanks Princeton University Observatory for hospitality and support and IAU commission 38 for a travel grant.

Received 19 March; accepted 14 May 1984.

- Gunn, J. E. *Ann. N.Y. Acad. Sci.* **375**, 287–296 (1981).
- Peacock, J. A. in *Quasars and Gravitational Lenses*, Proc. 24th Liege int. astrophys. Colloq. (ed. Swings, J. P.) 86–104 (University of Liege, 1983).
- Walsh, D., Carswell, R. F. & Weymann, R. J. *Nature* **279**, 381–384 (1979).
- Weymann, R. J. *et al. Nature* **285**, 641–643 (1980).
- Weedman, D. W., Weymann, R. J., Green, R. F. & Heckman, T. M. *Astrophys. J. Lett.* **255**, L5–L9 (1982).
- Lawrence, C. R. *et al. Science* **223**, 46–49 (1984).
- Djorgovsky, S. & Spinrad, H. *Bull. Am. astr. Soc.* **15**, 937 (1983).
- Young, P., Gunn, J. E., Kristian, J., Oke, J. B. & Westphal, J. A. *Astrophys. J.* **244**, 736–755 (1981).
- Turner, E. L., Ostriker, J. P. & Gott, J. R. *Astrophys. J.* (submitted).
- Bourassa, R. R. & Kantowski, R. *Astrophys. J.* **195**, 13–21 (1975).
- Gorenstein, M. V. *et al. Science* **219**, 54–56 (1983).
- Paczynski, B. & Gorski, K. *Astrophys. J. Lett.* **248**, L101–L104 (1981).
- Young, P., Deverill, R. S., Gunn, J. E. & Westphal, J. A. *Astrophys. J.* **244**, 721–735 (1981).
- Faber, S. M. & Jackson, R. E. *Astrophys. J.* **204**, 668–683 (1976).
- Schechter, P. L. & Gunn, J. E. *Astrophys. J.* **229**, 472–484 (1979).
- Felten, J. E. *Astr. J.* **82**, 861–878 (1977).
- Schneider, D. P., Gunn, J. E. & Hoessel, J. G. *Astrophys. J.* **264**, 337–355 (1983).
- Schechter, P. L. *Astrophys. J.* **203**, 297–306 (1976).
- Schmidt, M. & Green, R. F. *Astrophys. J.* **269**, 352–374 (1983).

- Dyer, C. C. & Roeder, R. C. *Astrophys. J.* **233**, L67–L70 (1980).
- Hoessel, J. G. *Astrophys. J.* **241**, 493–506 (1980).
- Burke, B. F., Roberts, D. H., Hewitt, J. N., Greenfield, P. E. & Dupree, A. K. in *Quasars and Gravitational Lenses*, Proc. 24th Liege int. astrophys. Colloq. (ed. Swings, J. P.) 203–207 (University of Liege, 1983).
- Narasimha, D., Subramanian, K. & Chitre, S. M. *Mon. Not. R. astr. Soc.* **200**, 941–950 (1982).
- Henry, P. Preprint, Univ. Hawaii.
- Shaver, P. A. & Robertson, J. G. *Astrophys. J. Lett.* **268**, L57–L61 (1983).
- Surdej, J., Swings, J. P., Henry, A., Arp, H., Kruszewski, A. & Pedersen, H. in *Quasars and Gravitational Lenses*, Proc. 24th Liege int. Astrophys Colloq. (ed. Swings, J. P.) 355–359 (University of Liege, 1983).
- Blandford, R. D. & Jaroszynski, M. *Astrophys. J.* **246**, 1–12 (1981).
- Bahcall, N. A. *Astrophys. J.* **198**, 249–254 (1975).
- Dressler, A. *Astrophys. J.* **226**, 55–69 (1978).
- Jones, C. & Forman, W. *Astrophys. J.* **276**, 38–55 (1984).
- Gott, J. R. & Turner, E. L. *Astrophys. J.* **216**, 357–371 (1977).
- Sol, H., Vanderriest, C., Lelievre, G., Pedersen, H. & Schneider, J. E. in *Quasars and Gravitational Lenses*, Proc. 24th Liege int. astrophys. Colloq. (ed. Swings, J. P.) 192–198 (University of Liege, 1983).
- Vanderriest, C. *et al.* in *Quasars and Gravitational Lenses*, Proc. 24th Liege int. astrophys. Colloq. (ed. Swings, J. P.) 182–191 (University of Liege, 1983).
- Keel, W. C. in *Quasars and Gravitational Lenses*, Proc. 24th Liege int. astrophys. Colloq. (ed. Swings, J. P.) 199–202 (University of Liege, 1983).
- Foltz, C. B., Weymann, R. J., Röser, H.-J. & Chaffee, F. H. *Astrophys. J. Lett.* **281**, L1–L4 (1984).

Light-inducible and chloroplast-associated expression of a chimaeric gene introduced into *Nicotiana tabacum* using a Ti plasmid vector

L. Herrera-Estrella, G. Van den Broeck, R. Maenhaut, M. Van Montagu & J. Schell

Laboratorium voor Genetica, Rijksuniversiteit Gent, B-9000 Gent, Belgium

M. Timko & A. Cashmore

Laboratory of Cell Biology, The Rockefeller University, New York, New York 10021, USA

A chimaeric gene, consisting of the 5'-flanking region of a member of the Pisum sativum gene family encoding ribulose 1,5-bisphosphate carboxylase linked to the coding region of a bacterial chloramphenicol acetyltransferase gene, has been introduced into the genome of the plant Nicotiana tabacum using a Ti plasmid of Agrobacterium tumefaciens. The expression of the chimaeric gene is light-inducible in chloroplast-containing transformed tissue.

USING the natural gene transfer system of *Agrobacterium tumefaciens*, it is now possible to introduce DNA sequences into plant cells. Chimaeric genes have been constructed in which bacterial coding sequences have been placed under the control of transcriptional signals of the nopaline synthase gene^{1–5}. When these chimaeric genes were introduced into plant cells using the Ti plasmid of *Agrobacterium*, the bacterial sequences were shown to be expressed as functional proteins^{1–5}. We have now used *Agrobacterium*-mediated transformation to study the regulatory sequences of a light-inducible plant gene.

The chloroplast enzyme ribulose 1,5-bisphosphate carboxylase/oxygenase is the major soluble protein of photosynthetic tissues and serves a bifunctional role in photosynthetic metabolism, catalysing both carboxylation and oxygenation of ribulose 1,5-bisphosphate. The carboxylase activity represents the first

step in the Calvin cycle and the oxygenase activity initiates the biosynthetic pathway of photorespiration. The enzyme is composed of eight large subunits of 53,000 molecular weight (MW) and eight small subunits of 14,000 MW. The large subunit protein is encoded by the chloroplast genome and synthesized on chloroplast ribosomes, whereas the small subunit protein is encoded by nuclear DNA and synthesized on free cytoplasmic ribosomes as a 20,000 MW precursor^{6–9}.

Studies in several plant species have shown that the small subunit protein is encoded by a small multi-gene family^{10–14}. Expression of these genes is regulated developmentally and in a tissue-specific manner. For example, the level of small subunit mRNA is much higher in plants grown in light than in dark¹⁵, and this increase is at least partially due to an increase in the level of transcription of the small subunit genes^{16,17}. In addition,

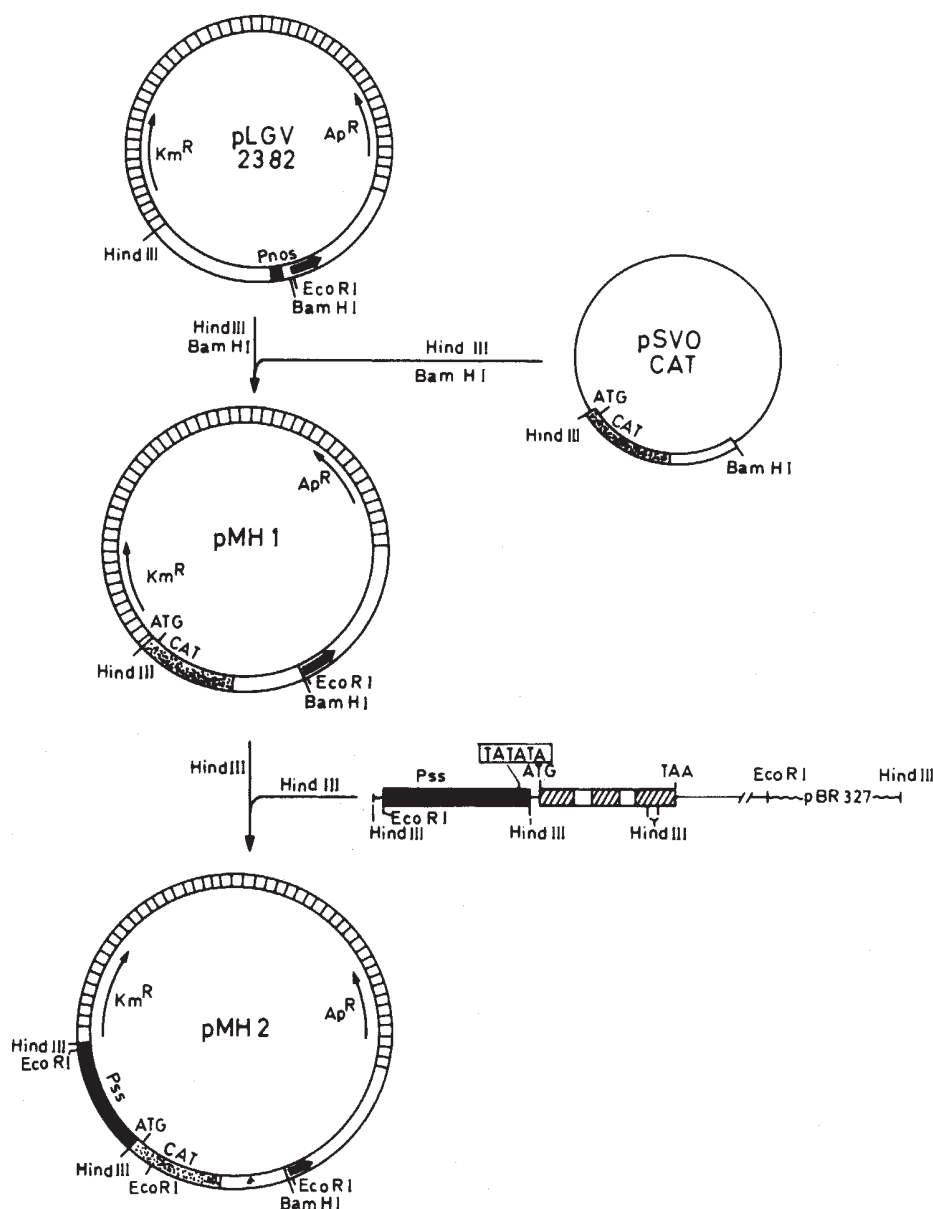


Fig. 1 Construction of a chimaeric gene containing the 5'-flanking sequence of the small subunit of ribulose-1,5-bisphosphate carboxylase linked to the chloramphenicol acetyltransferase (*CAT*) coding sequence. A 1.6-kb *HindIII*-*BamHI* fragment including the entire *CAT* coding sequence and the 3' end of the early SV40 transcript, was purified from a 1% agarose gel by electroelution. This fragment was used to replace the small *HindIII*-*BamHI* fragment of pLGV2382 by ligation at the *HindIII*-*BamHI* sites. The new vector (pMH1) contains, downstream of the *CAT* sequence, the functional signals from the *nos* gene for termination of transcription and polyadenylation. A *HindIII* fragment comprising 970 bp of the 5'-flanking sequences of the small subunit gene, starting 3 nucleotides upstream of the transcriptional start site, was cloned into the *HindIII* site of pMH1 to produce pMH2. ▨, pKC7 sequences³¹; →, 3' end region of the *nos* gene; ▤, *CAT* coding sequence; ■, 5'-flanking region of the small subunit gene (*Pss*); the coding region lies between *ATG* and *TAA*; ▤, introns; □, exons.

in both *Lemna* and *Pisum sativum* (pea), light induction is mediated by phytochrome¹⁸. In higher plants, genes for both the small and large subunit polypeptides are expressed in leaves but not in roots¹⁹, and in *C₄* plants both polypeptides are found in bundle sheath but not mesophyll cells of leaves^{13,20,21}. The expression of the nuclear genes encoding the small subunits is coordinately regulated, with the expression of the chloroplast genes encoding the large subunit^{16,22}. This regulation is particularly interesting as the copy number of the genes encoding these two polypeptides is vastly different, the genes for the large subunit being reiterated about a thousand-fold per cell in the numerous chloroplast genomes while the genes for the small subunit are present in only a few copies in the nuclear genome²³.

For animal genes, transcriptional regulation is often mediated by regulatory sequences residing 5' to the transcriptional initiation site^{24,25}. We have therefore explored the possibility that the small subunit gene has analogous 5'-transcriptional control signals which are responsible for its characteristic light-inducible expression. We report here the construction of a chimaeric gene (*ss-cat*) composed of the promoter of a *Pisum sativum* (pea) small subunit (*ss*) gene¹⁰ and the coding sequence of the bacterial gene for chloramphenicol acetyltransferase (*cat*). We have studied the expression of this chimaeric gene in tobacco cells.

Construction

To test whether sequences involved in the light-inducible expression of the small subunit gene are located in the region 5' to the start of transcription, a derivative of the vector pSVOCAT (ref. 26) was constructed. The *HindIII*-*BamHI* fragment from pSVOCAT containing the coding sequence of the *cat* gene and the polyadenylation site of the early transcript of simian virus (SV40), was substituted for the small *HindIII*-*BamHI* fragment of pLGV2382 (ref. 2) (Fig. 1). The new plasmid (pMH1) contains the 3' end of the nopaline synthase gene, thus providing a sequence which is known to be functional for both the termination of transcription and the polyadenylation of mRNA in plant cells¹. In addition, these sequences provide homology with the Ti plasmid used to introduce the chosen chimaeric gene into the T-DNA of the Ti plasmid by homologous recombination.

We have previously reported the sequence of a pea small subunit gene, *SS3.6* (ref. 10), and the cap site for this gene has now been determined by both primer extension and *S₁* nuclease studies (A.C., unpublished results). A conveniently located *HindIII* site lies four nucleotides 5' upstream of this cap site, which we presume corresponds to the point of initiation of



Fig. 2 The nucleotide sequence of the 5'-flanking region of the pea small subunit gene and partial sequence of the fusion point of the *ss-cat* chimaeric gene. The 5'-flanking sequence of the pea small subunit gene is presented starting with the first methionine codon (underlined) of the precursor polypeptide and extending 973 bp upstream from the start site for initiation of transcription. The nucleotide sequences are numbered by position +1 being the first nucleotide of the small subunit mRNA. The TATA box at position -31 is overlined. The lower case sequence represents pBR327 DNA, with the *EcoRI* site being the beginning of the pea genomic DNA. The sequence around the point of fusion of the pea small subunit untranscribed DNA (-4 to -973) and the CAT coding sequences are also depicted. The splice site was at *HindIII* restriction site noted in the sequence. The initiating methionine of the CAT polypeptide is underlined. The nucleotide sequences for the pea small subunit gene and the fusion point of the *ss-cat* chimaeric gene were determined by subcloning restriction fragments of the parent and derivative plasmids into M13 vectors and sequencing by the dideoxy method³².

transcription. We have now extended the sequence studies to include 973 nucleotides 5' from the cap site (Fig. 2). In order to construct a chimaeric gene in which the chloramphenicol acetyltransferase coding sequence was under the control of the small subunit promoter, the *HindIII* fragment containing the DNA sequences 4-973 nucleotides upstream of the transcriptional start of the small subunit gene (Fig. 2) was cloned into the *HindIII* site of pMH1 to produce pMH2 (Fig. 1).

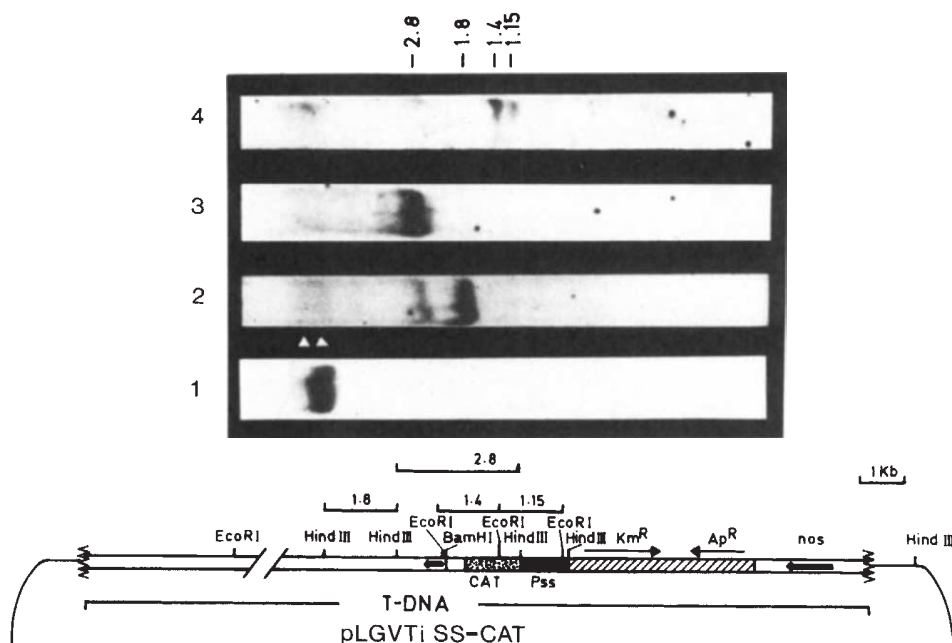
Transfer of the chimaeric gene

The *ss-cat* chimaeric gene was introduced into tobacco cells by first mobilizing pMH2 from *Escherichia coli* into *Agrobacterium* harbouring the wild-type pTiC58 plasmid. This was accomplished by using helper plasmids to provide in *trans* the Tra and Mob functions necessary for mobilization of pBR-derived plasmids²⁷. As pMH2 cannot replicate in *Agrobacterium*, kanamycin-resistant exconjugants can be obtained only when this plasmid co-integrates with the resident Ti plasmid. Such a co-integration occurs by a single recombination event through homology provided by the nopaline synthase sequence present in both plasmids. The presence of pMH2 sequences in the Ti plasmid was confirmed by Southern blot analysis²⁸ (data not shown), and the structure of one such co-integrate (pLGVTiSS-CAT) is shown in Fig. 3. *Agrobacterium* harbouring pLGVTiSS-

CAT was used to infect decapitated tobacco seedlings (Wisconsin 38, var. Petit Havana) and 2 weeks later the transformed tissue was transferred to Murashige and Skoog agar medium (without phytohormones) containing 500 mg l⁻¹ of cefotaximum (Claforan, Hoechst). The tissues were grown *in vitro* until axenic tissues were obtained. One randomly chosen tumour line (GVSSCAT) was used in all experiments described below.

To confirm the transfer and integration of the T-DNA carrying the *ss-cat* gene, genomic DNA prepared from tissue transformed with *Agrobacterium* harbouring pLGVTiSSCAT was analysed by Southern blot hybridization²⁸. The genomic DNA was digested with either the restriction enzymes *EcoRI* or *HindIII* and the fragments produced were separated by electrophoresis on a 1% agarose gel and transferred to nitrocellulose for hybridization analysis. Figure 3 shows the structure of the T-region in pLGVTiSSCAT and the results of the Southern hybridizations of genomic DNA from a transformed tissue chosen for further analysis. Using the coding sequence of the *cat* gene as a specific probe, two *EcoRI* fragments were detected (lane 4), a 1.15-kb fragment composed of part of the CAT sequence and 970 base pairs (bp) of the small subunits 5'-flanking region, and a 1.4-kb fragment derived from the 3' part of the *cat* gene and some sequences from pSVOCAT. Using the same probe, a 2.8-kb *HindIII* fragment was detected (lane 3)

Fig. 3 Structure of the T-region of pLGVTiSSCAT and Southern hybridization analysis of tobacco tissue containing the *ss-cat* gene. The lower part of the figure shows the structure of the T-region resulting from the co-integration of pMH2 with pTiC58 through the homology provided by sequences of the nopaline synthase gene present in both plasmids (←). The jagged lines represent the 25-bp sequence present as a direct repeat flanking the nopaline T-DNA. The upper part shows the results of the Southern hybridization analysis of genomic DNA isolated from tobacco cells transformed with pLGVTiSSCAT. Lanes 1 and 2 were digested with *Eco*RI and *Hind*III, respectively, and probed with a fragment containing homology to *Hind*III-23 and *Hind*III-31. Lanes 3 and 4 were digested with *Hind*III and *Eco*RI, respectively, and the filters hybridized with a probe specific for the CAT coding sequence. The numbers at the side of the autoradiogram represent the molecular weights of the most relevant fragments and are also indicated at the top of the diagram showing the T-region of pLGVTiSSCAT. All the symbols are as indicated in Fig. 1.



which corresponds to the entire CAT-coding sequence and the 3' end of the *nos* gene.

A probe specific for *Hind*III fragments 31 and 23 of the Ti plasmid was also used. The *Hind*III-23 fragment contains most of the sequences of the *nos* gene whereas fragment *Hind*III-31 is an internal fragment of the T-DNA which is adjacent and to the left of *Hind*III-23²⁹. This probe detected several *Hind*III fragments (Fig. 3, lane 2): a fragment of 2.8 kb corresponding to the same fragment detected with the CAT probe, a 1.8-kb fragment which corresponds to the T-DNA fragment *Hind*III-31, and two or more faint bands of high molecular weight which represent the junction fragments between the right end of the T-DNA and plant sequences. Because more than one band is detected with intensities lower than that found for the internal fragments of the T-DNA and because of data obtained from other Southern hybridization analysis (data not shown), we conclude that the two single copies of the T-DNA must have integrated into two different positions in the genome of the tobacco cells.

Light induction of *ss-cat*

The level of small subunit mRNA is strongly influenced by the lighting conditions in which plants are grown. These regulatory characteristics are also demonstrated in plant cell cultures, where small subunit genes are expressed in chlorophyll-containing tissue but are only weakly expressed in the absence of chlorophyll³⁰.

The chimaeric *ss-cat* gene provides a convenient enzymatic reporter for testing whether the 5'-flanking sequences derived from this particular small subunit gene from *Pisum sativum* can control transcription in tobacco cells. For this purpose, tobacco tissues containing the *ss-cat* gene were grown in different light conditions and tested for CAT activity. The results shown in the figures were all obtained with the tumour line GVSSCAT, but similar results were found with other tumour lines containing the *ss-cat* gene.

pLGVTiSSCAT-transformed tissue was first grown in Murashige and Skoog medium containing 0.5 mg l⁻¹ of cytokinin (benzylamino purine) and 30 mM sucrose, until leaf-like structures (teratoma) and mature chloroplasts were ob-

served. Subsequently, the tissue was transferred to medium without cytokinins and supplemented with sucrose to a final concentration of 3 mM. In these conditions, the semi-differentiated state of the tissues can be maintained stably for several weeks.

Chloroplast-containing tissue was divided into two equal portions and each piece was transferred to a separate plate containing the same medium. One sample was maintained in the dark while the other was kept in normal light conditions (photoperiod of 8 h dark and 16 h white light, 2,000 lx). After 5 days, the dark-grown tissue was divided into two, and one of the pieces transferred for 18 h to light. Equal amounts of the tissue, grown in normal light conditions, in the dark or in dark/light conditions, were ground and enzymatic CAT activity was determined for each of them as previously described¹.

Tissue grown continuously in normal light conditions showed a considerably higher level of CAT activity than tissue grown in the dark, in which only a low level of CAT activity was detected (Fig. 4). Similarly, tissue grown in the dark for 5 days and subsequently transferred to the light for 18 h showed a marked increase in the level of CAT activity. These experiments demonstrate that the *Pisum* small subunit promoter is functional in tobacco cells and suggest that it is responsible for the observed light-stimulated increase in CAT activity.

To test whether the expression of *ss-cat* is correlated with the presence of chloroplasts in the teratoma tissue, white unorganized tissue was tested for CAT activity after being grown for equal periods of time in normal light or dark conditions. Figure 4 (lane 5) shows that white tissue grown in normal light conditions contains a very low level of CAT activity and that this level does not vary considerably when the tissue is grown in the dark (Fig. 4, lane 6).

It was important to demonstrate that the observed light stimulation of CAT activity in tissue containing the *ss-cat* chimaeric gene is specific for the chimaeric gene and not due to a general and nonspecific increase in the level of transcription in light-treated cells. Therefore, as a control we used isogenic tobacco tissue containing another chimaeric gene, *nos-cat*¹, in which the CAT coding sequence is under the control of the constitutive promoter of the nopaline synthase gene.

Chloroplast-containing tissue with either the *ss-cat* or *nos-cat* gene was obtained, subcultured and maintained in light or darkness as described above. After 5 days the tissue was extracted and the enzymatic CAT activity determined. Figure 5 shows that the level of CAT activity in tissue containing *ss-cat* was strongly influenced by the light conditions in which the tissue was grown, whereas the level of CAT activity in tissue containing the *nos-cat* chimaeric gene was not affected by the light conditions. Quantitative analysis of the tests by counting the corresponding regions of the chromatograms in a liquid scintillation spectrometer showed that the expression of the *nos-cat* gene in dark-grown tissue was never less than 50% of that observed in light-grown tissue. CAT activity for the *ss-cat* gene in tissue grown in normal light conditions was 5–15-fold higher than the level present in tissue grown in the dark (Fig. 5). These results indicate that the light-inducible changes in enzymatic activity observed for the *ss-cat* gene are determined by the specific 5'-flanking promoter sequence.

A chimaeric gene construction containing the *ss* promoter in reverse orientation with respect to the *cat* coding sequence was used in similar control experiments. Transformed tissue harbouring this chimaeric gene showed no CAT activity above a very low background activity of nonspecific acetylation of chloramphenicol present in tobacco tissue transformed with *Agrobacterium* harbouring the wild-type Ti plasmid. This observation is consistent with the direction of transcription of the SS3.6 small subunit gene¹⁰.

To demonstrate that the observed induction of CAT activity was not due to variation in the overall rate of transcription or in the protein content of tissues grown in different light conditions, we used as internal control the levels of the nopaline synthase gene activity that is also present in the T-DNA of pLGVtISSCAT. Nopaline synthase activity in light-grown tissue was never found to be more than 50% higher than that present in dark-grown tissue; therefore, the increase in the level of CAT activity of light-grown tissue must reflect a differential increase in the rate of transcription of genes that are induced by light.

Measurement of mRNA

We wished to demonstrate that the observed changes in enzymatic activity in tissue harbouring the *ss-cat* gene are correlated with changes in the steady-state level of the corresponding poly(A)⁺ RNA. Poly(A)⁺ RNA was prepared from both green and white tissues grown in normal light conditions and from green tissue grown in the dark for 5 days. Figure 6 shows the results of dot-blot hybridization experiments involving different amounts of poly(A)⁺ RNA extracted from each tissue, and using as a probe a 773-bp *Taq*I fragment from Tn9 containing the CAT coding sequence. It can be observed that a significant level of RNA homologous to the CAT probe is present in differentiated tissue grown in the light. However, the level of this RNA is reduced ~10-fold when the same tissue is grown for 5 days in the dark. Also, a low but detectable level of transcription is found in white tissue containing the *ss-cat* gene. These results confirm that the light-induced changes in CAT enzymatic activity are indeed correlated with changes in the level of the corresponding poly(A)⁺ RNA.

Discussion

Experiments have been performed to examine the regulatory characteristics of a 5' promoter fragment of a pea small subunit gene. A chimaeric gene has been constructed in which this promoter fragment was linked to the coding sequences of the bacterial *cat* gene and to the 3' end of the *nos* gene. It was demonstrated that this gene was expressed on transference to tobacco cells and furthermore that the level of this expression was regulated by light conditions. An increase in the level of CAT enzymatic activity and also in the level of poly(A)⁺ RNA homologous to the *ss-cat* gene was observed in green tissue grown in the light when compared with the same tissue grown in the dark. In contrast, in non-differentiated white tissue only a low level of CAT activity was observed which did not increase

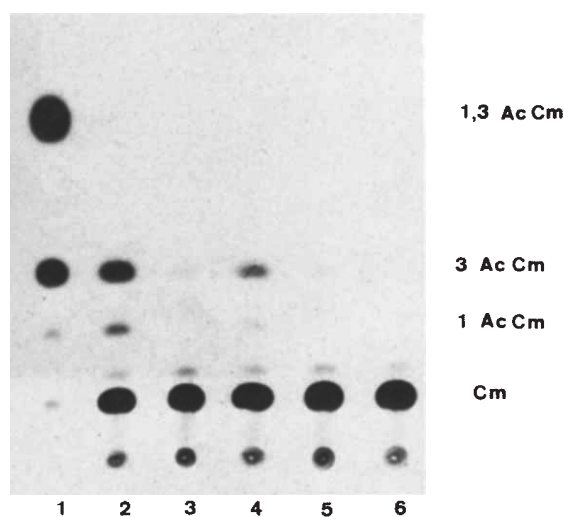


Fig. 4 Effect of the light on the level of CAT activity directed by the small subunit 5'-flanking sequences in tobacco tissues. Equal amounts of axenic tissues were ground and the CAT activity was determined as described previously. The reactions were allowed to proceed for 60 min at 37 °C, the samples were extracted with 1 ml of ethyl acetate, and analysed by ascending TLC using chloroform/methanol (95:5 v/v). X-ray films were exposed by the chromatograms and activities quantitated by counting the corresponding regions of the chromatograms in a liquid scintillation spectrometer. Lane 1, the CAT activity present in extracts of *E. coli* harbouring pBR325, as control. The following lanes show CAT activity present in extracts of tobacco tissue with the following light treatments: lane 2, chloroplast containing tissue grown in normal light conditions; lane 3, chloroplast containing tissue grown for 5 days in the dark; lane 4, chloroplast containing tissue grown for 5 days in the dark and 18 h under continuous light; lane 5, white tissue grown in normal light conditions; lane 6, white tissue grown for 5 days in the dark. ¹⁴C-chloramphenicol (Cm) and its acetylated derivatives (in order of increasing mobility), 1-acetylchloramphenicol (1 AcCm), 3-acetylchloramphenicol (3 AcCm) and 1,3-acetylchloramphenicol (1,3 AcCm) were detected by autoradiography.

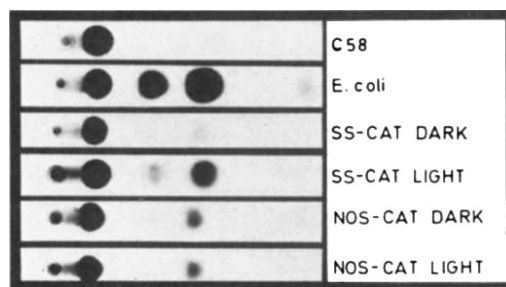


Fig. 5 Comparison of the effect of light on the expression of *ss-cat* and a constitutively expressed gene. CAT activity was determined in tobacco tissues as described in Fig. 4 legend. SS-CAT, green tobacco tissue containing the *ss-cat* gene grown for 5 days in the dark (DARK) or in normal light conditions (LIGHT). NOS-CAT, green tobacco tissue containing the *nos-cat* gene grown in the same dark or light conditions as the tissue containing the *ss-cat* gene. *E. coli*, CAT activity present in *E. coli*-harbouring pBR325. C58, tissue transformed with the wild-type plasmid, pTiC58, which does not contain any chimaeric gene. The percentage conversion of chloramphenicol to its acetylated derivatives by extracts containing 50 µg of protein and incubated for 30 min are as follows: *ss-cat* (dark), 1%; *ss-cat* (light), 17%; *nos-cat* (dark), 5.0%; *nos-cat* (light), 5.8%; C58, 0.3%. After long times of incubation and exposure of the respective film, a low nonspecific acetylation of chloramphenicol was observed. The different spots that can be observed in the autoradiogram correspond to chloramphenicol and its acetylated derivatives as shown in Fig. 4.

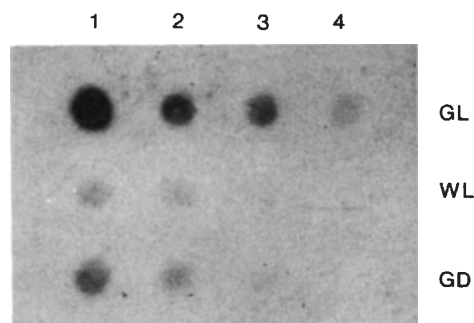


Fig. 6 Dot-blot analysis of poly(A)⁺ RNA homologous to the CAT sequences in tissue grown in different light conditions. Poly(A)⁺ RNA was extracted from tobacco tissue grown as described previously³³. Different amounts of RNA—5 µg (lane 1), 0.5 µg (lane 2), 0.1 µg (lane 3), 0.01 µg (lane 4)—were spotted on nitrocellulose filters with the use of the minifold device³⁴. A fragment containing the CAT coding sequence was labelled with ³²P by nick-translation (2×10^8 c.p.m. per µg) and used as a probe for hybridization with the RNA spots. GL, green tissue grown in normal light conditions; WL, white tissue grown in normal light conditions; GD, green tissue grown for 5 days in the dark.

when the tissue was grown in normal light conditions. It was also shown that these changes in expression are specific for the small subunit promoter, as the expression of a similar chimaeric gene using the *nos* promoter was independent of both light conditions and the presence of chloroplasts.

The experiments described are of interest for several reasons. Of primary interest is the demonstration that light induction is mediated by sequences residing 5' from the point of initiation of transcription of this pea small subunit gene. It is likely that the induction mechanism involves interaction between protein factors and this promoter fragment, and in view of this it is of interest that the pea small subunit promoter is able to mediate light-induction on transfer to tobacco cells. This observation must reflect significant conservation of the induction mechanism between these two species. The experiments are also useful for assessing the level at which this gene expression is regulated. We have shown that the expression of the *ss-cat* gene is light-inducible, in contrast to the *nos-cat* gene. As both chimaeric genes produce very similar transcripts which are novel for the plant cell, it is unlikely that they are subject to different post-transcriptional regulation. We conclude not only that the light-inducible expression of the *ss-cat* gene is mediated by the small subunit promoter sequences, but also that this induction must be regulated at the level of transcription initiation. Furthermore, as the chimaeric gene is expressed in chloroplast-containing callus tissue, we suggest that the tissue specificity of the light-inducible expression mediated by the *ss* promoter is associated with the ability of plant tissues to activate the assembly of the photosynthetic apparatus. This latter ability is probably determined by the levels of growth regulators (phytohormones) and

directly mediated by factors like phytochrome. Moreover, it has been shown that chlorophyll-containing roots (obtained in liquid culture containing cytokinins) also express the *ss* gene, while in normal roots the gene is not expressed at a detectable level (P. Eckes and L. Willmitzer, unpublished results).

The transformed tissues tested in these experiments contained more than one copy of the gene inserted at different positions in the tobacco genome. Because it is not known whether all copies are expressed and subject to regulation, further analysis is required to determine whether the site of insertion in the genome of the plant has any influence on light-regulated expression of the small subunit gene. We estimated that the level of poly(A)⁺ RNA specific for the CAT sequence in these transformed tobacco tissues is approximately 15 times lower than that found for the endogenous small subunit tobacco gene (data not shown). Many factors in addition to gene copy number may influence the level of expression of this chimaeric gene in tobacco cells. Whereas we have established that the pea small subunit promoter is clearly recognized and appropriately regulated by the transcription apparatus in tobacco cells, it may well be that this promoter does not function at the same level in tobacco as it does in pea. In addition, it is quite possible that regulatory sequences required for optimum expression exist elsewhere in the pea small subunit gene, in addition to those found within the 5'-flanking sequences. However, note that, in assessing the level of expression of *ss-cat* in tobacco cells, the particular small subunit gene used to provide the promoter in the chimaeric gene construction is only weakly expressed in pea leaves relative to other pea small subunit genes (A.C., unpublished results). To analyse these quantitative aspects of expression further, it is desirable to extend these studies to expression in leaf tissue. Using a pTi-derived vector which does not prevent normal differentiation of tobacco tissues⁵, we recently obtained tobacco plants containing the *ss-cat* gene. These plants express the *ss-cat* gene in their leaves and this expression was shown to be regulated by light conditions (unpublished results).

We thank Patricia Zambryski and Marc De Block for encouragement and advice; June Simpson, Ann Depicker, Patrick Dhase and Gilbert Engler for critical reading of the manuscript; Martine De Cock for preparation of the manuscript; and Albert Verstraete and Karel Spruyt for preparation of the figures. The research was supported by grants from the 'ASLK-Kankerfonds, the 'Instituut tot Aanmoediging van het Wetenschappelijk Onderzoek in Nijverheid en Landbouw (IWONL 3894A), the Services of the Prime Minister (OOA 12052179), the 'Fonds voor Geneeskundig Wetenschappelijk Onderzoek' (FGWO 3.001.82) to M.V.M. and J.S., the US Department of Health (GM31137) and the US Department of Energy (DE-AC02-80ER10581) to A.C., and was carried out under Research contract GVI-4-017-B (RS) of the Biomolecular Engineering Programme of the Commission of the European Communities. L.H.E. thanks CONACYT Mexico for a PhD fellowship, G.V.D.B. the Belgian IWONL for financial support, and M.T. the US Department of Health for a Research Service Award.

Received 27 March; accepted 22 May 1984.

- Herrera-Estrella, L., Depicker, A., Van Montagu, M. & Schell, J. *Nature* **303**, 209–213 (1983).
- Herrera-Estrella, L. *et al.* *EMBO J.* **2**, 987–995 (1983).
- Bevan, M. W., Flavell, R. B. & Chilton, M.-D. *Nature* **304**, 184–187 (1983).
- Fraley, R. T. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **80**, 4803–4807 (1983).
- Zambryski, P. *et al.* *EMBO J.* **2**, 2143–2150 (1983).
- Dobberstein, B., Blobel, G. & Chua, N.-H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1082–1085 (1977).
- Cashmore, A. R., Broadhurst, M. K. & Gray, R. E. *Proc. natn. Acad. Sci. U.S.A.* **75**, 655–659 (1978).
- Highfield, P. E. & Ellis, R. J. *Nature* **271**, 420–424 (1978).
- Chua, N. H. & Schmidt, G. W. *Proc. natn. Acad. Sci. U.S.A.* **75**, 6110–6114 (1978).
- Cashmore, A. R. in *Genetic Engineering of Plants—An Agricultural Perspective* (eds Kosuge, T., Meredith, C. P. & Hollaender, A.) 29–38 (Plenum, New York, 1983).
- Berry-Lowe, S. L., McKnight, T. D., Shah, D. M. & Meagher, R. B. *J. molec. appl. Genet.* **1**, 483–498 (1982).
- Coruzzi, G., Broglie, R., Cashmore, A. R. & Chua, N.-H. *J. biol. Chem.* **258**, 1399–1402 (1983).
- Broglie, R., Coruzzi, G., Lampka, G., Keith, B. & Chua, N.-H. *Bio/technology* **1**, 55–61 (1983).
- Dunsmuir, P., Smith, S. & Bedbrook, J. *Nucleic Acids Res.* **11**, 4177–4183 (1983).
- Bedbrook, J. R., Smith, S. M. & Ellis, R. J. *Nature* **287**, 692–697 (1980).

- Smith, S. M. & Ellis, R. J. *J. molec. appl. Genet.* **1**, 127–137 (1981).
- Gallagher, T. F. & Ellis, R. J. *EMBO J.* **1**, 1493–1498 (1982).
- Tobin, E. M. *Pl. molec. Biol.* **1**, 35–51 (1981).
- Thompson, W. F., Everett, M., Polans, N. O., Jorgensen, R. A. & Palmer, J. D. *Planta* **158**, 487–500 (1983).
- Link, G., Coen, D. M. & Bogorad, L. *Cell* **15**, 725–731 (1978).
- Broglie, R., Coruzzi, G. & Chua, N.-H. in *Chloroplast Biogenesis* (Cambridge University Press, in the press).
- Shinzaki, K., Sasaki, Y., Sakibama, T. & Kamikubo, T. *FEBS Lett.* **144**, 73–76 (1982).
- Cashmore, A. *Cell* **17**, 383–388 (1979).
- Brinster, R. L., Chen, H. Y., Warren, R., Sarthy, A. & Palmiter, R. D. *Nature* **296**, 39–42 (1982).
- Parker, M. *Nature* **304**, 687–688 (1983).
- Gorman, C. M., Moffat, L. F. & Howard, B. H. *Molec. cell. Biol.* **2**, 1044–1051 (1982).
- Van Haute, E. *et al.* *EMBO J.* **2**, 411–418 (1983).
- Southern, E. M. *J. molec. Biol.* **98**, 503–518 (1975).
- Depicker, A. *et al.* *Plasmid* **3**, 193–211 (1980).
- Vernet, T. *et al.* *Eur. J. Biochem.* **126**, 489–494 (1982).
- Rao, R. N. & Rogers, S. G. *Gene* **7**, 79–82 (1979).
- Sanger, F., Nicklen, S. & Coulson, A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
- De Greve, H. *et al.* *J. molec. appl. Genet.* **1**, 499–512 (1982).
- Thomas, P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201–5205 (1980).

Magnetic flare model of γ -ray bursts

E. P. Liang* & S. K. Antiochos†

* Lawrence Livermore National Laboratory, University of California, Livermore, California 94550, USA

† Institute for Plasma Research, Stanford University, via Crespi, Stanford, California 94305, USA

The thermal synchrotron (TS) interpretation of γ -ray burst continuum spectra¹ has recently gained support from the analysis of an expanding database²⁻⁵. However, this interpretation requires an emission region which is hot ($kT \sim 0.2-1mc^2$), optically very thin ($nh \leq 10^{21} \text{ cm}^{-2}$) with a highly super-Eddington flux $F \sim 10^{30} \text{ erg cm}^{-2} \text{ s}^{-1}$ ($\gg F_{\text{Edd}} = 10^{25} \text{ M/M}_\odot$) for a 10-km neutron star. This picture is similar to that first proposed for the 5 March 1979 event⁶⁻⁸. In addition there are hints that the emission layer is very dense ($n_e \leq 10^{24}-10^{26} \text{ cm}^{-3}$) and thin ($h \leq 10^{-3} \text{ cm}$). For example, events which show simultaneous redshifted 511-keV annihilation lines and low energy self-absorption (see Fig. 1) allow us to estimate a lower limit in the range $10^{23}-10^{26} \text{ cm}^{-1}$ (ref. 9) to the pair density, provided that the annihilation region coincides with the synchrotron source. To maintain the pair population close to Maxwellian, the collision excitation rate into higher Landau levels⁹ (that is, the pitch-angle scattering rate) must exceed the cyclotron decay rates. Coulomb scattering between pairs and protons is too slow. Even if collective processes whose rates are close to the electron plasma frequency, or scattering by heavy ions (for example, $Z = 26$, rates $\propto Z^2$), are invoked, a particle density $n \geq 10^{25} \text{ cm}^{-3}$ is still needed. Both arguments point towards a dense but thin emitting sheet with $h \leq 10^{-3}-10^{-5} \text{ cm}$. The severe energetics and persistence of such a hot, dense, thin emitting layer prompted us to consider a picture in which the emission regions lie at the foot points of reconnecting magnetic loops, powered by downward impinging electromagnetic waves. We now examine the structure of the emitting sheet, and the generation, propagation and coupling of the electromagnetic energy fluxes to the surface layer, and show that a flare-like model can account for most of the general features of γ -ray bursts.

If a thin plasma sheet located at the neutron star surface is the source of the γ rays, then the total energy emitted is too large, $\sim 10^{31-36} \text{ erg cm}^3$, to have been stored *in situ*; hence, it must originate in a much larger volume and be transported to the emitting region by an energy flux of some form. An external downward energy flux is preferred over an internal upward flux because the downward momentum flux automatically helps to hold down the plasma sheet. Otherwise, the sheet would expand along the field lines as it heated because the gravitational scale height is much larger than the above values for h . Nor can this external flux of energy be carried by a beam of particles (pairs or ions) as the density required would be too high to avoid violating the above limit on nh unless the beam were ultra-relativistic. But ultra-relativistic ($E \gg mc^2$) particles would be expected to deposit their energy over a column depth $\gg 10^{21} \text{ cm}^{-2}$. We therefore conclude that the energy is both stored and transported primarily in the form of electromagnetic (EM) energy. This suggests a burst model analogous to that proposed for solar flare X-ray bursts, where the energy source is believed to be stressed magnetic field in the corona and the hard X-ray emission originates from footpoints of reconnecting flux tubes¹⁰ (Fig. 2) (see refs 11, 12 for previous discussions of the flare idea).

As well as satisfying the severe constraint on nh , there are other attractive features to the flare model. Assuming a field strength of $3 \times 10^{12} \text{ G}$, sufficient energy can be stored in a neutron star volume, $E \sim 10^{42} R_6^3 \text{ ergs}$, to power all the observed galactic bursts ($d \leq 20 \text{ kpc}$). Note that the magnetic energy available to power the burst is only the non-potential component due to currents in the corona, that is the magnetic free energy. Assuming

that the currents are due to particles drifting at relativistic velocities, a coronal density $N \sim 10^{15} B_{12}/R_6 \text{ cm}^{-3}$ is obtained. This value is low enough to satisfy the restriction on nh . In addition, the effective coronal density may be lower if, as expected, the currents are concentrated into thin filaments¹³.

Another encouraging feature of this model is that it can account for the two distinct time scales in γ -ray bursts: $\sim$ milliseconds for the risetime of individual spikes and $\sim$ seconds for the total burst duration. Although the detailed process by which the magnetic free energy is released is unknown, even for solar flares, two physical time scales are important. One is the Alfvén transit time which for our case is given by: $\tau_a \sim 10^{-4} R_6 \text{ s}$, as the Alfvén speed is limited by c for the large field strengths and low densities in our model. This time scale can account for even the most rapid time structure in γ -ray bursts. The other time scale of interest is the resistive tearing time¹⁴ given by: $\tau \sim (\tau_a \tau_d)^{1/2}$ where τ_d is the resistive diffusion time scale. We expect the plasma to be in a highly turbulent state and the resistivity to be greatly enhanced above collisional values. Assuming an effective collision frequency due to the collective process of the order of the plasma frequency yields: $\tau \sim 10(N_{15})^4 (R_6)^{3/2} \text{ s}$, which can account for the burst duration. Numerical simulations have shown that reconnection can proceed through a burst process¹⁵, which results in two distinct time scales.

Because the coronal density is so low we expect that the released magnetic free energy will not be absorbed directly by the coronal plasma. In the classical picture for magnetic annihilation, such as the Petschek model¹⁶, energy initially in a d.c. magnetic field is converted to an energy flux that propagates away from a small reconnection region and is convected along the background field (Fig. 2). This energy flux can be considered as a general form of magnetohydrodynamic (MHD) wave flux because both plasma and magnetic field will be convected away from the reconnection region. For subrelativistic propagation velocities, $v/c \ll 1$, we expect approximate equality between the plasma kinetic energy, $\rho(v)^2$, and the a.c. magnetic energy, $(\delta B)^2/8\pi$, in these 'waves'. For relativistic propagation velocities, $v/c \approx 1$, and for our extremely low densities, the plasma kinetic energy will be negligible, $\rho c^2 < (\delta B)^2/8\pi$. Instead there will be approximate equipartition between the a.c. magnetic energy and the a.c. electric energy, $\delta E \approx \delta B v/c \approx \delta B_1$ as is the case for relativistic Alfvén waves¹⁷.

An important question is whether such waves can couple effectively to the neutron star surface. We anticipate that non-linear effects may cause the bulk of the power to cascade down to wavelengths of the order of the shortest physical scale, the ion Larmor radius. In that case the dominant frequency would be $\nu_0 \sim 10^{16} B_{12}^{-1} \text{ s}^{-1}$ which is comparable to the plasma

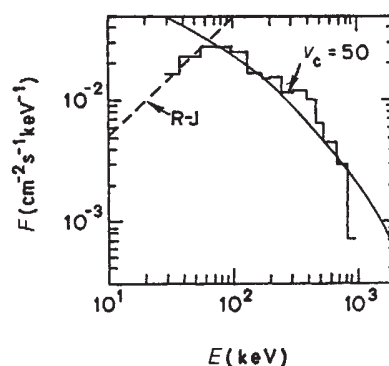


Fig. 1 Spectrum of a γ -ray burst on 4 February 1979, showing evidence for an emission line at $\sim 400 \text{ keV}$ and a Rayleigh-Jeans spectrum at low energies, $\leq 50 \text{ keV}$ (ref. 1).

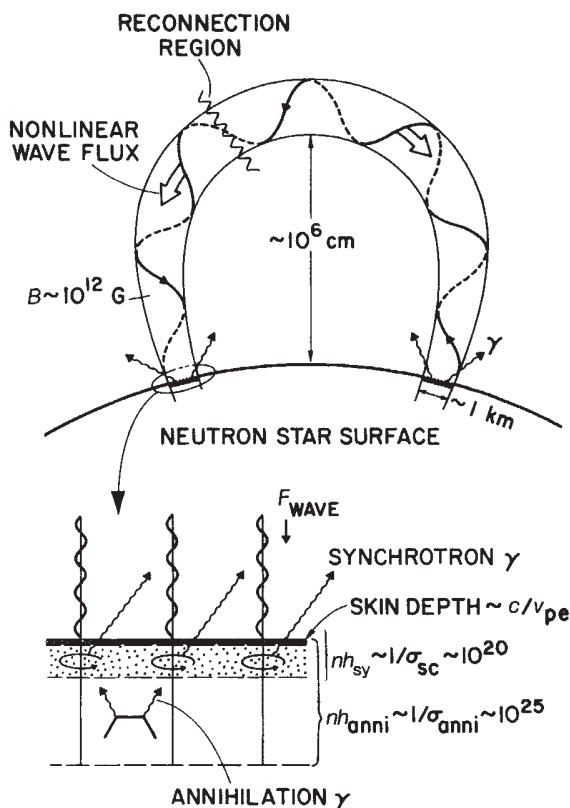


Fig. 2 Conceptual visualization of a flaring flux tube on a neutron star with a γ -ray emission layer at each of the loop footpoints on the stellar surface.

frequency of the γ -ray emitting region, $\nu_{pe} \sim 10^{16} (n_{25})^{1/2} \text{ s}^{-1}$. Because of the ultrahigh flux and steep density profile at the stellar surface, we believe that the waves will be absorbed mainly through nonlinear collective plasma processes, such as resonant absorption, Raman scattering, parametric decay and so on, similar to the situation in laser-target interactions¹⁸. These nonlinear plasma processes convert wave energy efficiently and collisionlessly into hot suprathermal electrons with characteristic temperature T_{hot} up to tens to hundreds of keV, on the local electron plasma oscillation time scale ($\sim 10^{-16} \text{ s}$). In our case, these hot electrons (pairs) would be the source of the synchrotron radiation as they would probably be born with large momenta transverse to the ambient B -field ($\delta E \perp B_0$). As in the case of laser-target interaction, the density profile of the background electrons would be steepened by the impinging wave momentum flux, creating an extremely thin acceleration region with a coherent electrostatic acceleration field, of the order of the Debye length¹⁸ ($\sim$ incident wavelength $< 10^{-6} \text{ cm}$). The hot electrons, after leaving this region, could travel only a short distance before cooling off. The thickness of the hot synchrotron layer is then roughly the synchrotron cooling distance $\nu_{\text{hot}} t_{\text{syn}} \sim 10^{-5} \text{ cm}$, consistent with the observed column depth of $nh \sim 10^{20} - 10^{21} \text{ cm}^{-2}$. Depending on which mode of nonlinear plasma collisionless process is operating, the hot electron temperature would have well-defined scaling properties with the incident wave flux¹⁸ ($\propto$ emergent γ flux). Such scaling behaviour could, in principle, be checked by future observations.

Finally, let us consider the ultimate source of the burst energy, the mechanism by which the neutron-star magnetic field is sheared. There are several possibilities: one is that there is a process similar to that on the Sun, in which motions due to either differential rotation or crust or core restructuring stress the coronal field. Recent observations of the 5 March event may have discovered a 23-ms quasi-periodicity that may be related to torsional vibrations of a neutron star¹⁹, suggesting that shearing movements of a neutron star surface are not impossible.

Another possibility is that the field is stressed by accreting material. Note that in our context the accreted gravitational energy is not released directly. First it is transferred to the field in the form of magnetic stress, which then results in EM energy propagating along the field to the star surface. Hence, the mass that provides the burst energy is likely to be far away from the γ -ray emitting region, and provides no opacity to the γ rays. It is conceivable that surface nuclear flashing of accreted material may also stress the field²⁰. But this would occur at rather large optical depths. It is difficult to see how it can produce an optically-thin γ -ray burst instead of an optically-thick X-ray burst.

The Lawrence Livermore National Laboratory is operated under DOE contract W-7405-Eng-48; the Institute for Plasma Research at Stanford is partially supported by NASA grant NGR-05-020-668.

Received 23 January; accepted 26 April 1984.

1. Liang, E. P., Jernigan, T. E. & Rodrigues, R. *Astrophys. J.* **271**, 766-777 (1983).
2. Mazets, E. P. et al. *Astrophys. Space Sci.* **80**, 3-83 (1981).
3. Barat, C. *Proc. Santa Cruz Workshop on High Energy Transients* (ed. Woosley, S.) (AIP, New York, 1984).
4. Desai, U. D. et al. *Proc. COSPAR Cosmic-Ray Symp.*, Bangalore (Publisher? 1983).
5. Nolan, P. L. *Proc. Santa Cruz Workshop on High Energy Transients* (ed. Woosley, S.) (AIP, New York, 1984).
6. Ramaty, R., Lingenfelter, R. & Bussard, R. W. *Astrophys. Space Sci.* **75**, 193-203 (1981).
7. Ramaty, R. et al. *Nature* **287**, 122-124 (1980).
8. Katz, J. I. *Astrophys. J.* **260**, 370-385 (1982).
9. Liang, E. P. *AIP Conf. Proc. No. 101* (eds Burns, M., Harding, A. & Ramaty, R.) 81-88 (AIP, New York, 1983).
10. Duijveman, A., Hooyng, P. & Machado, M. E. *Sol. Phys.* **81**, 137-157 (1982).
11. Ruderman, M. *Ann. N.Y. Acad. Sci.* **262**, 164 (1975).
12. Katz, J. I. in *AIP Conf. Proc. No. 101* (eds Burns, M. L., Harding, A. K. & Ramaty, R.) 65-75 (AIP, New York, 1983).
13. Parker, E. N. *Astrophys. J.* **264**, 635-641 (1983).
14. Furth, H. P., Killeen, J. & Rosenbluth, M. N. *Phys. Fluids* **6**, 459-484 (1963).
15. Priest, E. R. *AGU Chapman Conf. on Reconnection*, Los Alamos (American Geophysical Union, Washington DC, 1983).
16. Petschek, H. E. *AAS-NASA Symp. on the Physics of Solar Flares*, 425-430 (NASA Spec. Publ. SP-50, 1964).
17. Barnes, A. & Suffolk, G. C. J. *J. Plasma Phys.* **5**, 315-329 (1971).
18. Max, C. in *Les Houches Summer School, Proc. Laser-Plasma Interactions*, (eds Balian, R. & Adam, J. C.) (North-Holland, Amsterdam, 1981).
19. Barat, C. et al. *Astr. Astrophys.* **126**, 400-402 (1983).
20. Woosley, S. in *AIP Conf. Proc. No. 77* (ed. Lingenfelter, R.) 273-292 (AIP, New York, 1982).

Detection of new molecules in the visible spectrum of Comet IRAS-Araki-Alcock (1983 d)

Cristiano B. Cosmovici

Space Research Group, DFVLR, 8031 Wessling, FRG

Sergio Ortolani

Asiago Astrophysical Observatory, 36012 Asiago (VI), Italy

The discovery of Comet IRAS-Araki-Alcock provided the first opportunity for studying a comet at a distance of $4.5 \times 10^6 \text{ km}$ from the Earth, thus allowing a spatial resolution of $\sim 20 \text{ km}$ to be achieved. Photographic and spectroscopic plates obtained on 9.9 UT May 1983 using the 90-cm Schmidt and the 182-cm Copernicus telescopes at the Asiago Observatory have revealed new fundamental discoveries. The photometric data system (PDS)-elaborated Schmidt plates show that the inner coma of the comet is directed towards the Sun with an inclination of $\sim 30^\circ$, a phenomenon probably due to the rotation of the nucleus. On the spectroscopic plates taken with the grating spectrograph in the spectral range 3,800-8,500 Å with a resolution of 6 Å, 450 lines can be distinguished, although 90 lines are still unidentified. After detailed analysis, we have identified the molecules HCO and H_2S^+ which have never before been observed in the visible part of the spectrum of any astronomical object. We suspect that H_2CO , DCO, S_2 and NH_4 are also present. We report here only the preliminary results from our spectroscopic data.

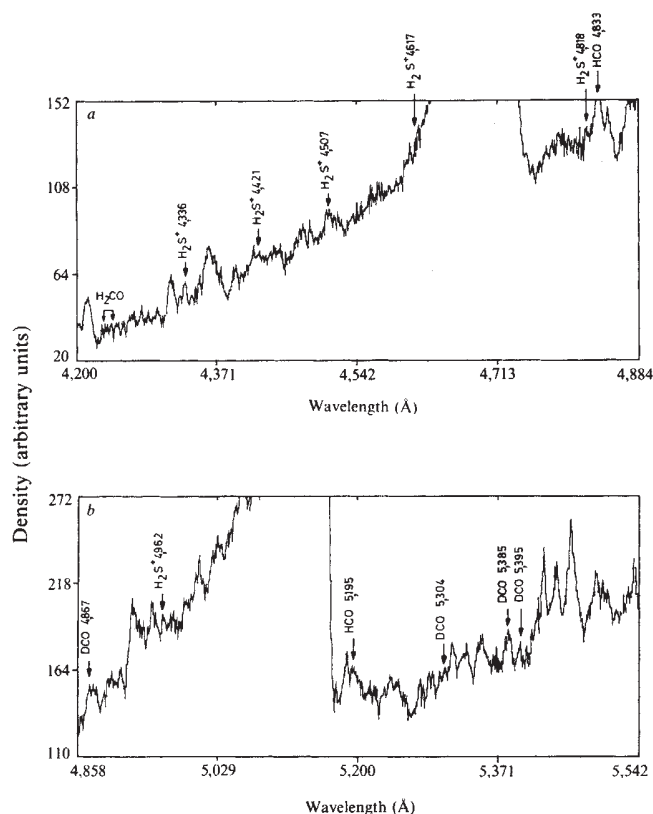
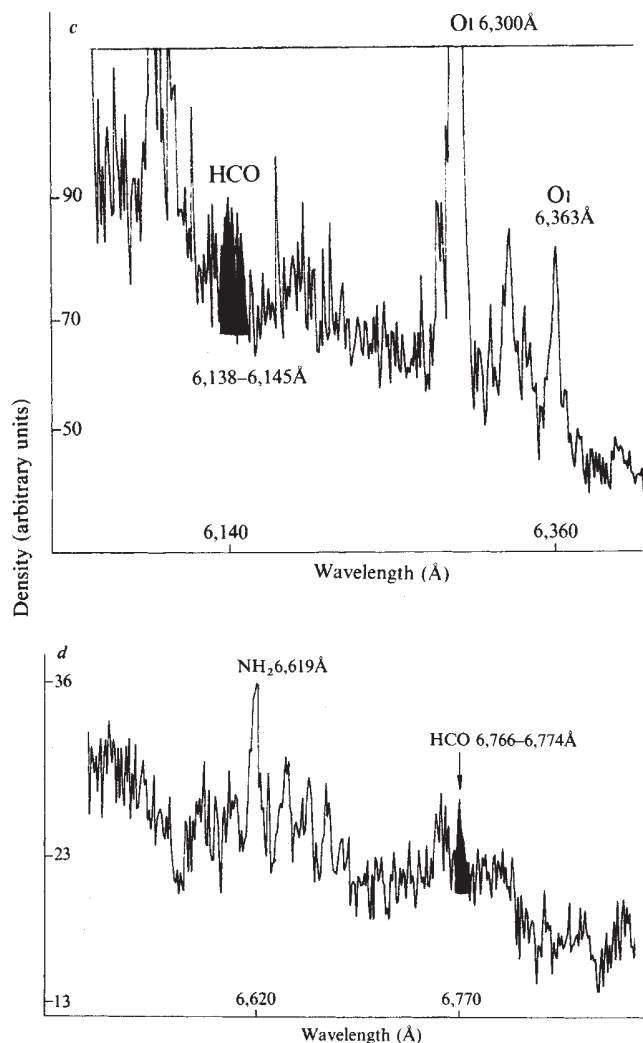


Fig. 1 Four regions of spectrum no. 2922 taken at 9.9 UT May 1983 with the BC 600 + Varo grating spectrograph at the Asiago 182 cm Copernicus Telescope and processed with the PDS at ESO (Garching) with a resolution of 6 Å. Slit 10,000 × 150 km centred on the brightest part of the coma (north-south direction). Exposure time 30 min on 103-a-D plate. The flux of the HCO-band at 4,833–38 Å is of the order of $5 \times 10^{-11} \text{ erg cm}^{-2} \text{ s}^{-1}$.

During the observation, the comet was 0.052 AU from the Earth and 1.01 AU from the Sun. The phase angle was 87.7° and the angle of the vector from the Sun to the comet was 171.3° . The spectra were taken on 103-a-D Kodak plates with exposure time varying from 5 to 30 min with a BC 600 + Varo grating spectrograph. The measurements were done at exactly the same time as those of the IRAS satellite¹, thus comparison of the IR and visible data is extremely important. We have no notification of visual observations on 9.9 May from other observers.

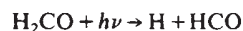
S₂ was discovered in the UV by the IUE satellite on 11 May (ref. 2), and H₂O, NH₃ by the 100-m Effelsberg radio telescope on 11 and 12 May³. The slit of our spectrograph covers a region of 10,000 × 150 km centred on the brightest part of the inner coma towards the Sun. The four spectra (see Fig. 1) were analysed with the PDS at the European Southern Observatory (ESO), Garching: all known neutral and ion molecules were found to be present. This analysis confirmed the discovery of the triplet and Asundi bands of CO (refs 4–6) and enabled the first identification of the CO-Å system (0–5, 0–4, 0–3, 0–1, 0–0, 1–1, 1–0); this system will be discussed elsewhere.

The strongest unidentified features at 4,833–4,838, 5,159–5,201, 6,138–6,145 and 6,766–6,774 Å belong to the formyl radical (HCO), which has already been discovered in interstellar clouds at 3.46 mm (ref. 7). The other strong laboratory lines at 5,624–5,629 and 5,148–5,153 Å merge with very strong C₂ bands. The Herzberg group⁸ has only observed absorption spectra of HCO in the laboratory. It is difficult to observe this molecule in emission at vibrational levels up to $v_2 = 15$ because of the

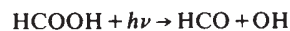


strong predissociation occurring at intermediate levels (G. Herzberg, personal communication). Thus we have been wary of publishing our discovery, but recent laboratory measurements enabled the laser-induced fluorescence detection of HCO ($v_2 = 9$ and $v_2 = 13$) produced by the laser photolysis of formaldehyde⁹.

Figure 1 shows the observed HCO lines in our best spectrum (the Q and R branch cannot be clearly separated at our resolution). The line at 4,833–4,838 is anomalously strong with respect to the other HCO lines⁸; but this could be due to contamination of CO (Ångström system at 4,835 Å)¹⁰. The HCO flux at 4,838 Å is of the order of $5 \times 10^{-11} \text{ erg cm}^{-2} \text{ s}^{-1}$. The most probable production mechanism of HCO is from the photodissociation of formaldehyde



It is also possible that the formic acid detected in interstellar clouds at 18 cm (see ref. 11) could dissociate as:



It would be very difficult for HCO to be the parent in the frozen nucleus because it is a primary intermediate in hydrocarbon oxidation. Otherwise, HCO is very unstable and has a dissociation energy (D) of only $\sim 1 \text{ eV}$ ($D_{\text{H}_2\text{CO}} = 3.37 \text{ eV}$).

Figure 2 shows that HCO disappears at $\sim 1,000 \text{ km}$ from the nucleus in the direction of the Sun, but a second maximum at 670 km represents a region where HCO is produced—possibly by dissociation of H₂CO.

Table 1 New molecules in Comet IRAS (1983 d)

Identification positive	(ν'_1, ν'_2, ν'_3)	$\lambda(\text{\AA})$	Ref.
HCO	(0, 7, 0)	6,766–6,774	8, 10
	(0, 9, 0)	6,138–6,145	
	(0, 11, 0)	5,624–5,629	
		(blended by C_2)	
	(0, 13, 0)	5,195–5,201	
	(1, 9, 0)	5,148–5,152	10
		(blended by C_2)	
	(0, 15, 0)	4,833–4,838	
H ₂ S ⁺		4,962, 4,818, 4,755, 4,692 (blended by C_2), 4,617, 4,507, 4,421, 4,336, 4,012 (blended by C_3)	10
Strongly suspected H ₂ CO		4,121, 4,220–40, 4,434, 4,551–69, 4,695–73, 4,821, 4,347 (blended by CH + NH ⁺), 4,044 (blended by C_3), 3,952 (blended by C_3 + CH ⁺ + CO ⁺)	
S ₂		Main system: 4,433, 4,651, 4,791, 4,937, 5,710, 5,841 (all other laboratory lines blended) Red system: 6,984, 7,070, 7,319	
DCO	(0, 19, 0)	4,867–4,870	8
	(0, 17, 0)	5,385–5,389	
	(0, 13, 0)	5,842–5,847	
	(0, 5, 0)	8,123–8,130	
		(other bands blended):	
	(0, 15, 0)	5,472–5,476	Herzberg
	(0, 11, 0)	6,273–6,278	
	(0, 9, 0)	6,780–6,785	
		Diffuse bands: 5,304, 5,395, 6,057, 5,490, 5,670, 5,696, 6,322	
NH ₄		6,628–6,635	

HCO and H₂CO were discovered in interstellar clouds by radio-astronomy^{7,11}, S₂ in the IRAS comet by the IUE satellite², DCO only in ionic form in interstellar clouds¹¹ and H₂S only as a neutral in interstellar clouds¹¹. A strong line at 4,837.64–4,838.39 was discovered in Comet Mrkos (1957 d) and given as unidentified¹³, although Greenstein and Arpigny¹⁴ suspected that this line belongs to HCO, but had no evidence for other lines. Larson published¹⁵ evidence for S₂ in the visible but no other data have followed.

The low dissociation energy of HCO makes the detection of this molecule possible only when a very high spatial resolution close to the Earth can be achieved. The mass spectrometer on board the Giotto spacecraft may enable HCO and H₂CO to be detected, but because of the volatility of these molecules the abundance in a periodic comet such as Halley could be insignificant if they are formed only in the outer shell of the nucleus.

Processes like the Alfvén mechanism^{5,6} could be responsible for the excitation of HCO, but further theoretical investigations are necessary to understand the excitation mechanisms in that particular case.

We have pointed out⁴ that the lines of H₂S⁺ at 4,421, 4,509 and 4,756 Å could be present in Comet Bradfield 1980 t, but because of the bad resolution and the low intensity of these lines we could not confirm the discovery. In the case of Comet IRAS the strong line at 4,507 Å is very clear and all laboratory known strong lines are observable (see Table 1). H₂S ($D = 3.77$ eV, $I_p = 10.42$ eV) has been observed at 1.8 mm in interstel-

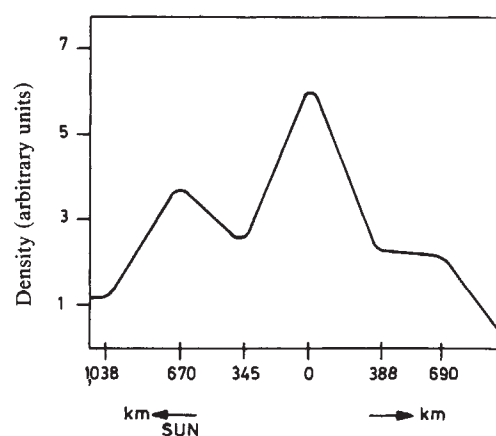


Fig. 2 PDS-profile of the HCO-line at 4,833 Å, obtained directly from the TRACE command of the Image Handling and Processing system without any smoothing or filtering procedure. Only the continuum subtraction has been made on the original scan. The smooth appearance can be explained by the fact that the PDS scan aperture slit used was $100 \times 5 \mu\text{m}$ with x and y steps of the same corresponding size. The smoothing is the result of the y size of the slit, which is much higher than the typical noise frequency of the photographic emulsion grain. On the abscissa 0 km represents the centre of the nuclear brightest region. The second peak at 670 km is probably due to dissociation of H₂CO (or HCOOH).

lar clouds⁴, but we now have the first cosmic evidence for H₂S⁺. Studies are being carried out to determine the spatial distribution of H₂S⁺.

For the other suspected new molecules reported in Table 1, further investigations are necessary. P. Feldman and M. A'Hearn (personal communication) report possible evidence of H₂CO UV-lines in their IUE-spectra and theoretical studies by Ip¹² support the detectability of molecules containing deuterium in cometary nuclei. Herzberg reports that the strong line at 6,633 Å has been seen by Spinrad in Comet Bradfield 1979 I. This feature is actually a doublet with a spacing of 3 Å and coincides with the Schüler band of NH₄.

Comet IRAS (1983 d) may have been the only opportunity for spectroscopic studies at high spatial resolution and for detecting new fundamental molecules by means of ground-based observations. The space missions to Comet Halley and the space observatories' Space Telescope and Infrared Space Observatory might confirm and improve these discoveries within the next 10 yr.

We thank Professors G. Herzberg, L. Biermann, F. Whipple, M. A'Hearn, P. Feldman, W. Hübner, C. Arpigny, and Drs M. Festou and H. Porsche for valuable discussions and suggestions, also U. Ritter for help with spectra analysis.

Received 28 February; accepted 9 May 1984.

1. Walker, R. G. *et al. Astrophys. J. Lett.* **278**, L11 (1984).
2. A'Hearn, M. F., Feldman, P. D. & Schleicher, D. G. *Astrophys. J. Lett.* **274**, L176 (1983).
3. Altenhoff, W. J. *et al. Astr. Astrophys.* **125**, L19 (1983).
4. Cosmovici, C. B., Barbieri, C., Bonoli, C., Bortolotto, F. & Hamzaoglu, E. *Astr. Astrophys.* **114**, 373 (1982).
5. Cosmovici, C. B., Biermann, L. & Arpigny, C. *Proc. ESO Workshop, The Need for Coordinated Ground-based Observations of Halley's Comet*, 131–147 (ESO, 1982).
6. Arpigny, C., Biermann, L. & Cosmovici, C. B. *Proc. int. Conf. on Cometary Exploration*, Budapest Vol. 1, 185 (Hungarian Academy of Sciences, Budapest, 1983).
7. Hollis, J. M. & Churchwell, E. *Astrophys. J.* **271**, 170 (1983).
8. Johns, J. W. C., Priddle, S. H. & Ramsay, D. A. *Discuss. Faraday Soc.* **35**, 90 (1963).
9. König, R. & Lademann, J. *Chem. Phys. Lett.* **94**, 152 (1983).
10. Pearse, R. W. B. & Gaydon, A. G. *The Identification of Molecular Spectra* 4th edn (Chapman & Hall, London, 1976).
11. Cosmovici, C. B., Inguscio, M., Strafella, F. & Strumia, F. *Astrophys. Space Sci.* **60**, 475 (1979).
12. Ip, W. H. *Proc. NATO Workshop on Ices in the Solar System*, Nice (Reidel, Dordrecht, 1984).
13. Woszczyk, A. *Mém. Soc. R. Sci. Liège, Fasc.* **6** (1962).
14. Greenstein, J. C. & Arpigny, C. *Astrophys. J.* **132**, 892 (1962).
15. Larson, S. M. *IAU Circ. No.* 3811 (1983).

A new interstellar molecule: tricarbon monoxide

Henry E. Matthews*, William M. Irvine†, Per Friberg‡§, Ronald D. Brown‡ & Peter D. Godfrey‡

* Herzberg Institute for Astrophysics, National Research Council of Canada, Ottawa, Ontario, Canada K1A 0R6

† Astronomy, GRC Tower B, University of Massachusetts, Amherst, Massachusetts 01003, USA

‡ Department of Chemistry, Monash University, Clayton, Victoria, Australia 3168

The cold dark interstellar Taurus Molecular Cloud One (TMC-1) is a rich source of acetylenic and polyacetylenic molecular species. As well as linear closed-shell molecules ($\text{H}(\text{C}\equiv\text{C})_n\text{CN}$) and symmetric rotors ($\text{CH}_3\text{C}\equiv\text{CH}$, $\text{CH}_3\text{C}\equiv\text{CCN}$), several radicals ($\text{C}\equiv\text{CH}$, $\text{C}\equiv\text{CCN}$, $(\text{C}\equiv\text{C})_2\text{H}$) have also been identified¹⁻⁵, many of which had not been studied previously in the laboratory. Whether the observed abundances can be understood in terms of purely gas-phase ion-molecule chemical schemes, which produce reasonable agreement for the simplest polyatomic species, is unclear^{6,7}; alternative models involving the particulate interstellar grains as catalysts or sources have also been suggested^{8,9}. We now report the detection in TMC-1 of a new molecule, tricarbon monoxide (C_3O), whose pure rotational spectrum has only very recently been studied in the laboratory. As C_3O is the first known interstellar carbon chain molecule to contain oxygen, its existence places an important new constraint on chemical schemes for cold interstellar clouds. In fact, the observed abundance of tricarbon monoxide fits quite well into our model of galactochemistry.

The observation of six microwave lines, together with accurate *ab initio* molecular orbital calculations, observed Stark patterns, and an appropriate mass spectrometer peak, conclusively demonstrated the presence in the laboratory of a new gas-phase linear molecule with an electronic structure well represented by the classical resonance formula $\text{C}\equiv\text{C}-\text{C}\equiv\text{O}^+$ (ref. 10). Our astronomical detection was carried out using a frequency-switching technique at the National Radio Astronomy Observatory's 43-m telescope in Green Bank, West Virginia. The half-power beamwidth was 1.5 arc min, and the aperture and beam efficiencies were, respectively, 0.27 and 0.36. Figure 1 shows the spectrum obtained towards the position of maximum cyanopolyne emission in TMC-1 ($\alpha(1950) = 4\text{ h } 38\text{ min } 38.6\text{ s}$; $\delta(1950) = 25^\circ 35' 45''$) at a spectral resolution of 12 kHz = 0.18 km s^{-1} . The velocity scale assumes the measured rest frequency $\nu_0 = 19,243.531 \pm 0.020\text{ MHz}$ for the $J = 2-1$ transition of C_3O .

Assuming the velocity (v) observed for a variety of other related molecules in this source ($v = 5.8-5.9\text{ km s}^{-1}$)^{1-5,11}, the astronomical emission line coincides with the laboratory measurement to within our spectral resolution, or 1 p.p.m. On the basis of local oscillator shifts carried out during our observations, we can rule out instrumental and terrestrial sources for the feature. As proof of the assignment to C_3O , we note the subsequent detection at the same location in TMC-1 of the $J = 9-8$, $8-7$ and $5-4$ transitions, which will be discussed elsewhere¹².

The line in Fig. 1 has a small velocity width ($\sim 0.2\text{ km s}^{-1}$), as we confirmed by obtaining higher resolution (0.09 km s^{-1}) data immediately following its initial detection. In this respect it is similar to several other species, such as C_4H (ref. 11) and like them C_3O thus seems to be confined chiefly to one of the two cloud components at this position towards TMC-1.

We made sensitive searches for the $J = 2-1$ transition of C_3O towards other objects, some of which are rich molecular-line sources, and which together encompass a wide range of physical and (presumably) chemical conditions. We were unable to detect C_3O towards any of these objects. Our upper limits are given in Table 1, in terms of the statistical noise of the spectra.

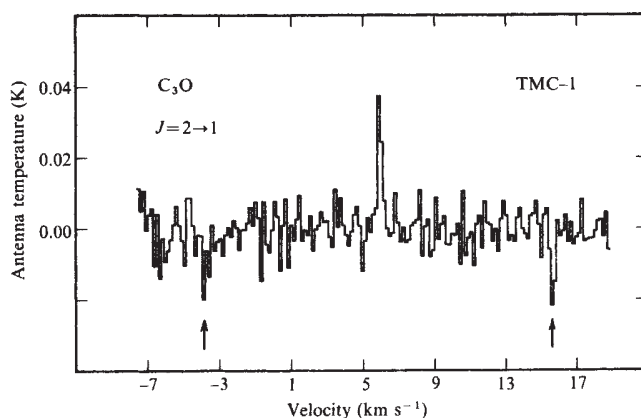


Fig. 1 A spectrum of the dust cloud TMC-1 showing the $J = 2-1$ transition of C_3O . The velocity scale is correct for a rest frequency of 19,243.531 MHz, and the ordinate is antenna temperature corrected for atmospheric absorption and the telescope gain as a function of hour angle. Negative half-intensity images of the line are expected at the positions marked by arrows, as a result of the frequency-switching technique used.

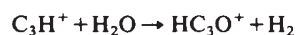
Table 1 Sources searched for $\text{C}_3\text{O}(J = 2-1)$

Source	RA (1950) h min s	Dec (1950)	V_{LSR}^* (km s^{-1})	Resolution (km s^{-1})	$3 \times \text{r.m.s.}$ ($\text{mK } T_{\text{a}}^*$)
W3(OH)	02 23 16.5	61° 38' 57"	-44.0	0.74	49
DR21(OH)	20 37 14.0	42 12 00	-3.0	0.97	14
Orion KL	05 32 47.0	-05 24 20	9.0	0.74	23
W49N	19 07 52.5	09 01 17	8.0	2.92	17
W51M	19 21 26.2	14 24 43	62.0	3.90	8
Sgr B2	17 44 11.3	-28 22 30	62.0	3.90	15
L134N	15 51 32.0	-02 43 00	2.0	0.18	13
TMC-1	04 38 38.6	25 35 45	5.8	0.18	See Fig. 1
IRC +10 216	09 45 14.8	13 30 39	-26.0	7.8	2
Cas A	23 20 55.0	58 32 18	0.0	0.97	12

* Assumed velocity with respect to local standard of rest.

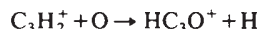
The rotational temperature for C_3O in TMC-1 almost certainly falls in the range $3 \leq T_{\text{rot}} < 10\text{ K}$, where the limits are set by the cosmic black-body background (the line is obviously in emission) and the kinetic temperature determined for this cloud in several studies¹³. In fact, comparison with the excitation for other species with similar molecular weight and electric dipole moment ($\mu(\text{C}_3\text{O}) = 2.39\text{ D}$)^{10,11} suggests $T_{\text{rot}} \approx 5-10\text{ K}$. Adopting this range of values and assuming that the line is optically thin, we calculate a column density $N(\text{C}_3\text{O}) \approx 1(10)^{12}\text{ cm}^{-2}$. This is about two orders of magnitude less than that of the C_4H radical, and roughly one-seventh the abundance of the radical C_3N (refs 11, 13, 14). Although with less certainty, an abundance by number relative to molecular hydrogen can be obtained by taking the ratio of the above column density with that of CO (actually to $^{12}\text{C}^{18}\text{O}$) and adopting the value $[\text{CO}]/[\text{H}_2] = 8(10)^{-5}$ determined for dark clouds¹⁴⁻¹⁶; this yields $[\text{C}_3\text{O}]/[\text{H}_2] \approx 10^{-10}$.

The detection and determination of the fractional abundance of C_3O allows us to test some hypotheses about the mechanisms of formation of long-chain carbon compounds, one of the more puzzling current aspects of galactochemistry. Significantly, C_3O does not contain nitrogen, which indicates that the latter element is not crucial in carbon chain formation. Tricarbon monoxide is plausibly formed by electron recombination with the ion HC_3O^+ . Our chemical model of molecular clouds (refs 17, 18, and R.D.B. and E. Rice, unpublished data), extended to incorporate *inter alia* the hydrocarbon chain-building reactions described by Schiff and Bohme¹⁹, suggests that the most important formation channels for HC_3O^+ are

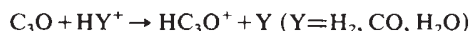
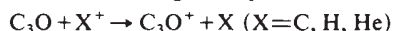


§ Permanent address: Onsala Space Observatory, S-43900 Onsala, Sweden.

with an additional production channel being



We note that C_3H^+ and C_3H_2^+ (but not C_3H_3^+) are expected constituents of TMC-1, according to recent models and laboratory reaction data^{6,7}. The important destruction channels for C_3O are expected to be charge and proton transfer reactions:



Using plausible physical parameters for TMC-1 (kinetic temperature 10 K, H_2 density 10^4 cm^{-3}), our model shows a C_3O

abundance relative to molecular hydrogen that varies with cloud age, reaching a peak of 10^{-7} at 10^6 yr and a relatively steady value $\sim 3(10)^{-11}$ at 10^8 yr. The latter is in good agreement with the observational result suggesting that, at least in cold clouds such as TMC-1, the abundance of C_3O can be understood in terms of purely gas-phase ion-molecule chemistry.

The National Radio Astronomy Observatory is operated by Associated Universities, Inc., under contract with the NSF. This research was supported in part by NASA grant NAGW-436 and by the Australian Research Grants Scheme. P.F. acknowledges a fellowship from the Swedish Natural Sciences Research Council (NFR).

Received 8 February; accepted 10 May 1984.

1. Avery, L. W. in *Interstellar Molecules* (ed. Andrew, B. H.) 47–58 (Reidel, Dordrecht, 1980).
2. Irvine, W. M., Höglund, B., Friberg, P., Askne, J. & Ellender, J. *Astrophys. J. Lett.* **248**, L113–L117 (1981).
3. Broten, N. W. *et al. Astrophys. J. Lett.* **276**, L25–L29 (1984).
4. Wootten, A., Bozayan, E. P., Garrett, D. B., Loren, R. B. & Snell, R. L. *Astrophys. J.* **239**, 844–854 (1980).
5. Friberg, P., Hjalmarson, Å., Irvine, W. M. & Guélin, M. *Astrophys. J. Lett.* **241**, L99–L103 (1980).
6. Herbst, E., Adams, N. G. & Smith, D. *Astrophys. J.* **269**, 329–333 (1983).
7. Millar, T. J. & Freeman, A. *Mon. Not. R. astr. Soc.* (in the press).
8. Hayatsu, R. & Anders, E. *Topics Curr. Chem.* **99**, 1–37 (1981).

9. Sagan, C. & Khare, B. *Nature* **277**, 102–107 (1979).
10. Brown, R. D., Eastwood, F. W., Elmes, P. S. & Godfrey, P. D. *J. Am. chem. Soc.* **105**, 6496–6497 (1983).
11. Guélin, M., Friberg, P. & Mezaoui, A. *Astr. Astrophys.* **109**, 23–31 (1982).
12. Brown, R. D. *et al.* (in preparation).
13. Irvine, W. M. & Hjalmarson, Å. in *Cosmochemistry and the Origin of Life* (ed. Ponnamperuma, C.) 113–142 (Reidel, Dordrecht, 1983).
14. Irvine, W. M. & Hjalmarson, Å. *Origins of Life* (in the press).
15. Frerking, M. A., Langer, W. D. & Wilson, R. W. *Astrophys. J.* **262**, 590–605 (1982).
16. Guélin, M., Langer, W. D. & Wilson, R. W. *Astr. Astrophys.* **107**, 107–127 (1982).
17. Brown, R. D. *Chem. N. Z.* **42**, 58–62 (1978).
18. Brown, R. D. & Rice, E. *Phil. Trans. R. Soc. A303*, 523–533 (1981).
19. Schiff, H. I. & Bohme, D. K. *Astrophys. J.* **232**, 740–746 (1979).

Micrometre-sized polymer layers synthesized by MeV ions impinging on frozen methane

G. Foti, L. Calcagno & K. L. Sheng*

Istituto Dipartimentale di Fisica, Corso Italia 57, I-95124 Catania, Italy

G. Strazzulla

Osservatorio Astrofisico, Città Universitaria, I-95125 Catania, Italy

There has recently been much experimental and theoretical interest among astrophysicists in the synthesis of complex organic materials (polymers) as a consequence of the energy deposition by external agents on carbon-containing molecular solids such as CH_4 , C_2H_6 , C_3H_8 , C_6H_6 and mixtures with CO_2 , CO , NH_3 and H_2O . Solid residues have been obtained by bombardment of hydrocarbon rich frozen layers with UV photons¹ and energetic particles^{2–4}. We present here the first quantitative results on the rate of production of solid residues, obtained by bombarding thick ($\approx 10^{19}$ – 10^{20} C atoms cm^{-2}) frozen layers ($T \approx 4$ K) of methane with 1.5-MeV protons produced at the Catania University van de Graaff accelerator. This process converts the methane molecules to a polymer-like residue which is stable at room temperature and above. The thickness of the solid residues has been measured both *in situ* by elastic proton scattering (in C atoms cm^{-2}) and remotely, by talysurf measurements (in μm). Thicknesses of up to 15 μm or 3.1×10^{19} C atoms cm^{-2} have been measured and the density of the residues is $\rho \geq 0.5 \text{ g cm}^{-3}$. The process appears to occur along the entire path of the incoming protons.

Thick films of methane were accreted by vapour deposition on a cold finger in contact with a liquid helium Dewar placed in the scattering chamber. The vacuum in the chamber was maintained at better than 10^{-7} mbar to ensure that only layers of background gases consisting of at most a few molecules can be added to the thick frozen layer of methane. The temperature of the finger was not measured directly, but film stability was tested by measuring its thickness at various times after deposition. The methane solids were then bombarded with a 1.5-MeV

proton beam at various fluences (10^{15} – 5×10^{16} ions cm^{-2}) and currents (1–100 nA cm^{-2}).

To study the part played by the substrate, for example, through its thermal conductivity, different materials were mounted on the finger before methane deposition: (1) a $\sim 1\text{-}\mu\text{m}$ gold film on a silicon wafer, (2) a silicon wafer polished like a mirror, (3) potassium chloride cleaved from a monocrystal and immediately put in the chamber and (4) a $3\text{-}\mu\text{m}$ H_2O ice layer. Material (4) was deposited on (1) and (2) just before the methane was deposited.

Film thickness was measured before and after bombardment using the elastic scattering of the 1.5-MeV proton beam. A typical energy spectrum of scattered protons is given in Fig. 1b: the dashed area is proportional to the number of C atoms cm^{-2} in the target⁵. The thickness of the films as deposited (T_0) ranges from 0.65×10^{19} to 3.9×10^{19} C atoms cm^{-2} , corresponding to 3 and 19 μm for frozen methane, which has a density⁶ of 0.555 g cm^{-3} .

During bombardment at low temperature, the thickness of the film was recorded by the same incoming beam to test whether T_0 changed drastically due to sputtering or evaporation. However, the decrease in T_0 was not detectable because for 1.5-MeV protons the sensitivity is not high enough to detect low erosion yields. The lowest number of C atoms detectable by proton scattering in the geometry of our experiment is $\sim 5 \times 10^{17} \text{ cm}^{-2}$ because the signal which comes from carbon overlaps the high background signal. Sputtering measurements performed with higher sensitivity for the 1.5-MeV He^+ beam give an erosion yield of 120 C atoms per He^+ (ref. 7 and W. L. Brown and L. J. Lanzerotti, personal communication); thus we expect a yield for MeV H^+ on methane of the order of 3 C atoms per H^+ (ref. 8). With this value and a proton dose of 5×10^{16} ions cm^{-2} , we obtain a total erosion of 10^{17} C atoms cm^{-2} that is not detectable by our technique.

After bombardment, the sample was warmed and the thickness of the solid residue measured at room temperature. The experimental data obtained for two different thicknesses of deposited films are plotted against proton fluence in Fig. 1a. The data can be fitted using the simple relationship:

$$T = T_0 (1 - e^{-\sigma\phi}) \quad (1)$$

where T (C atoms cm^{-2}) is the thickness of the residue, T_0 (C atoms cm^{-2}) is the thickness of the frozen methane as deposited, σ (cm^2) is the cross-section of the target area, ϕ (protons cm^{-2}) is the total implanted dose. Equation (1) is shown as the solid curves in Fig. 1a. From these we find $\sigma = 4 \times 10^{-17} \text{ cm}^2$ for 1.5-MeV H^+ on solid methane. The exponential dependence of

* Permanent address: Institute of Nuclear Research, Academia Sinica, PO Box 8204, Shanghai, China.

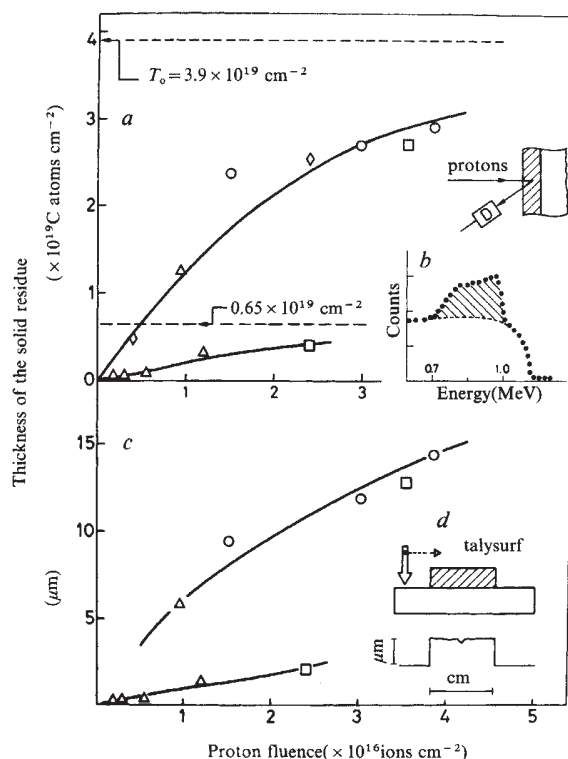


Fig. 1 *a*, Thickness of the synthesized residues measured by proton scattering plotted against proton dose for two initial thicknesses and various substrates: $\circ$, silicon wafer; $\diamond$, KCl; $\square$, ice; $\triangle$, gold. *b*, Energy spectrum of protons scattered from the 2.7×10^{19} C atoms cm^{-2} residue on the silicon substrate. *c*, Thickness of the synthesized residues measured by talysurf (*d*) plotted against proton fluence for the two initial thicknesses given in *a* and various substrates (key as *a*).

the final thickness of the solid residue layer proves that the phenomenon investigated is a bulk process, where the methane molecules act as targets⁹. Along the ion track, the high-energy deposition rate ($\sim 50 \text{ keV } \mu\text{m}^{-1}$ for 1.5-MeV H^+) produces, in a region of $\sim 50 \text{ \AA}$ around the ion path, a high concentration of broken bonds. The recombination of the new unstable species containing carbon atoms as CH_n radicals, gives rise to long-chain stable molecules mainly consisting of CH_2 and CH_3 groups, revealed by IR analysis of residues¹⁰. Radicals such as H have a high probability of recombining with each other and escaping from the target as H_2 with consequent hydrogen depletion; this was monitored by mass spectrometry *in situ* with the scattering chamber (refs 2, 11 and W. L. Brown and L. J. Lanzerotti, personal communication).

After bombardment, the targets were extracted from the scattering chamber and the thickness of the residues was measured by talysurf (Fig. 1*d*). The talysurf technique tests the physical profile of a sample surface with a moving mechanical finger. The sensitivity of the method is $\sim 500 \text{ \AA}$ for a sharp film boundary on a mirror-like surface, such as a silicon wafer used in micro-electronic devices. The surface of the residue appears as a yellow-brown spot of $5 \times 5 \text{ mm}$ and occasionally exhibits a rough profile with valleys and hills having lateral dimensions of the order of 0.2 mm .

The physical thicknesses are plotted in Fig. 1*c* against proton fluence for the same films given in Fig. 1*a*, with the exception of samples on KCl, for which talysurf fails because of the rough KCl surface. The relationship between thickness and fluence is very similar to that obtained with proton scattering. However, equation (1) cannot be applied. Of course, knowing the density of the methane layer as deposited and its thickness in C atoms cm^{-2} , we could calculate the initial thickness in μm . This, however, is not exactly comparable with the thickness of the

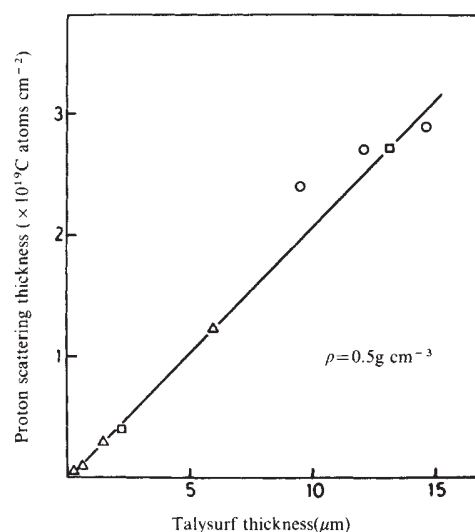


Fig. 2 Thickness of residues in C atoms cm^{-2} plotted against the same but in μm . The density of the material is deduced assuming a stoichiometry of C:H = 1:2.5.

solid residue because of changes in density and in the stoichiometry (ratio C/H) induced by the beam fluence. The solid curves in Fig. 1*c* are then drawn only as a guide.

The maximum thickness of the residue depends on the initial value and, for $T_0 = 3.9 \times 10^{19}$ C atoms cm^{-2} , reaches a value of $15 \mu\text{m}$. By increasing T_0 , the maximum thickness of the residue also increases. However, an upper limit is set by the range of the bombarding ion, which for a 1.5-MeV proton in methane is $\sim 40 \mu\text{m}$. In principle, by using more penetrating beams or on a continual removal of deposited methane layers, it should be possible to build up very thick layers. Of course, if the beam energy changes, the cross-section of the process would also change significantly. The dependence of σ on the stopping power of the incoming ion has still to be studied in detail.

Figure 2 plots two independent sets of thickness measurements for our residues against each other, and their correlation is quite good. Assuming, according to IR measurements¹¹, that our residues have a stoichiometry C:H = 1:2.5, we obtain a density for the new synthesized organic materials of $\rho = 0.5 \text{ g cm}^{-3}$. This value has to be considered as the lower limit because microfractures and voids in the film have been observed. Polymers such as polyethylene or polystyrene exhibit densities¹² between 0.9 and 1.2 g cm^{-3} ; however, lower values¹² between 0.3 and 0.6 g cm^{-3} have been measured in polyacetylene films depending on how the material is processed. More work is necessary to clarify the microscopic structure of the synthesized films.

We believe the present results have primary astrophysical significance. Frozen methane has been recognized as a major constituent of the surfaces of Pluto and Triton¹³. The presence of large amounts of hydrocarbons on icy mantles in interstellar dust and comets is also widely accepted¹. All these objects are continuously bombarded by energetic ions from cosmic rays or solar wind and solar flares that could be an important role in the build up of new organic materials, similar to those synthesized in our laboratory, both on interstellar grains⁴ and on planetary objects¹¹. The finding that IR features observed from our residues are the same as from some astrophysical objects¹⁰ supports this belief.

Received 16 January; accepted 3 May 1984.

- Greenberg, M. J. in *Comets* (ed. Wilkening, L. L.) 131–163 (University of Arizona Press, 1982).
- Daube, M. E., Augustiniak, W. M., Brown, W. L., Lanzerotti, L. J. & Poate, J. M. *EOS* **59**, 314 (1978).
- Moore, M. H. & Donn, B. *Astrophys. J. Lett.* **257**, L47–L50 (1982).
- Strazzulla, G., Calcagno, L. & Foti, G. *Mon. Not. R. astr. Soc.* **204**, 59p–62p (1983).

5. Mayer, J. M. & Rimini, E. (eds) *Ion Beam Handbook for Material Analysis* (Academic, New York, 1977).
6. *Handbook of Chemistry and Physics* 62nd edn (CRC Press, Boca Raton, 1981–82).
7. Johnson, R. E., Lanzerotti, L. J., Brown, W. L., Augustiniak, W. M. & Mussil, C. *Astr. Astrophys.* **123**, 343–346 (1983).
8. Lanzerotti, L. J., Brown, W. L., Poate, J. M. & Augustiniak, W. M. *Nature* **272**, 431–433 (1978).
9. Calcagno, L., Strazzulla, G., Fichera, M. & Foti, G. *Rad. Eff. Lett.* **76**, 143–148 (1983).
10. Strazzulla, G., Cataliotti, R. S., Calcagno, L. & Foti, G. *Astr. Astrophys.* (in the press).
11. Strazzulla, G. *Icarus* (in the press).
12. Mori, D. J. & Pfister, G. *Electronic Properties of Polymers* (Wiley, New York, 1982).
13. Apt, J., Carleton, N. P. & Mackay, C. D. *Astrophys. J.* **270**, 342 (1983).

Possible role of cloudiness in the persistence of the Southern Oscillation

Peter B. Wright

Joint Institute for the Study of Atmosphere and Ocean,
University of Washington, Seattle,
Washington 98195, USA

The Southern Oscillation (SO) exhibits persistence which can be explained by positive feedback between ocean and atmosphere¹. The persistence exhibits an annual cycle, being weakest in the northern winter and spring. Analysis of a marine data set shows that sea surface temperature (SST) in the equatorial east Pacific is positively correlated with cloudiness in the northern winter but not in the summer and autumn. Because increased cloudiness favours a reduction in net incoming radiation, the relationship implies a reduced positive feedback on the SST in winter. I show here that this seasonal variation of the cloudiness feedback could be the main cause of the seasonal variation in persistence.

Monthly mean SST data from ship reports, averaged over 2-degree squares, were obtained from J. O. Fletcher (personal communication), and converted into anomalies in 4° lat. by 10° long. boxes relative to the mean for the same calendar month in 1949–68. The mean SST anomaly over the region of the east Pacific outlined in Fig. 1, a convenient index of the SO (ref. 2), exhibits strong persistence which varies markedly with season (Fig. 2). The interannual variability, represented by the standard deviation (s.d.) of the index, also changes markedly from a minimum in March to a maximum in November (Fig. 2). Data are for 1947–69. The same index for 1900–39, although less reliable because there were fewer observations, exhibited a very similar pattern (not shown). As a convenient measure of their similarity, the correlation between the column for lag 3 in Fig. 2 and the corresponding column for 1900–39 (C_3) is 0.88; the 1% significance level for 12 independent values is 0.71. Further, the index is highly correlated with a rainfall index of the SO (ref. 2) which may be regarded as an inverse measure of the strength of the Walker circulation, and which exhibits similar patterns of both persistence and variability¹ (C_3 between the indices is 0.95).

I have shown previously that the above behaviour could be accounted for on the hypothesis that there is a feedback relationship between the SST and the Walker circulation in which the feedback varies from strongly positive about October to weakly positive about April. This was demonstrated using a simple feedback model of the form

$$\frac{dA}{dt} = -k_a A + k_a q B + k_a R_a \quad (1)$$

$$\frac{dB}{dt} = -k_b B + k_b p A + k_b R_b \quad (2)$$

where A represents the negative of the Walker circulation and B the east Pacific SST, k_a , k_b are positive constants representing the decay rates of anomalies, q and p represent the interactions between A and B , and R_a , R_b are first-order Markov processes. The observed statistics could be simulated if the product pq is always positive but varies from 0.4 to 1.3 in different seasons.

Qualitative physical reasoning^{1,6} suggests that q is always positive (warm sea favours weaker Walker circulation). Several mechanisms contribute to p . Weaker Walker circulation is associated with weaker upwelling and advection which favour warming of the sea, so the contribution of these mechanisms to p is positive. Weaker Walker circulation is also generally associated with more cloudiness^{3,4}, which in the equatorial zone causes a reduction in net incoming radiation⁵ and therefore favours cooling of the sea, so this mechanism makes a negative contribution to p . Let

$$p = p_0 + p_1 C \quad (3)$$

where p_0 is the net contribution of all mechanisms other than cloudiness and is assumed positive, C is a measure of the association between Walker circulation and cloudiness, and p_1 is negative to represent the assumption that more cloud favours colder sea. I now present evidence that C varies seasonally, and that these variations, through their contribution to p , could be the main cause of the seasonal variations displayed in Fig. 2.

Cloudiness data were obtained from the same source as the SST data and were similarly processed. Figure 1 shows regression coefficients of monthly anomalies of cloudiness on the SST index for the period 1947–69. Over the year as a whole, warm sea is associated with a marked positive anomaly of cloudiness in the equatorial central Pacific, but in other regions the association with cloudiness is small and of either sign. These results agree well with those presented by Egger *et al.*³ based on satellite data for 1965–73 for September to February, and those of Liebmann and Hartmann⁴ based on satellite data for individual seasons during 1974–78. These encouraging similarities between patterns based on different data sets for different periods support the reality of the relationships and indicate that the present cloudiness data set is useful for such analyses.

Table 1, based on cloudiness averaged over three regions (Fig. 1, dashed lines), confirms a strong positive correlation in the western region, a weak positive in the central region, and a negligible relationship in the east. Table 2 shows the regression coefficients for the same data stratified by month. In the western sector (column 1) the correlation is positive throughout the year, whereas in the eastern sector (column 3) it is positive from December to February and negative in most other months. The

Table 1 Relations between cloudiness in the three sectors in Fig. 1 and the SST index of the Southern Oscillation, all months 1947–69

Sector	180–150° W	150–120° W	120–90° W
Regression (oktas per °C)	0.35	0.13	–0.04
Correlation	0.47	0.19	–0.03
No. months data	209	182	276

Table 2 Regression coefficients of cloudiness in the four regions outlined in Fig. 1 on the SST index of the Southern Oscillation

	1 180–150° W	2 150–120° W	3 120–90° W	4 Region of SST index
January	63	49	17	32
February	58	50	21	40
March	45	10	–13	2
April	17	9	–2	6
May	41	20	1	15
June	34	27	–16	4
July	31	12	–11	0
August	25	–13	1	2
September	34	–1	–23	–4
October	39	16	–14	0
November	15	–1	–14	–2
December	33	–5	15	12

Units, oktas $\times 100$ (°C)^{–1}.

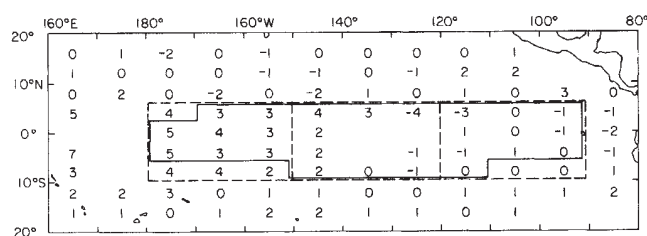


Fig. 1 Regression coefficients of cloudiness on SST index, all months 1947-69. Units, oktas $\times 10$ ($^{\circ}\text{C}$) $^{-1}$. Solid line, region used to define the SST index. Dashed lines, regions used in analyses.

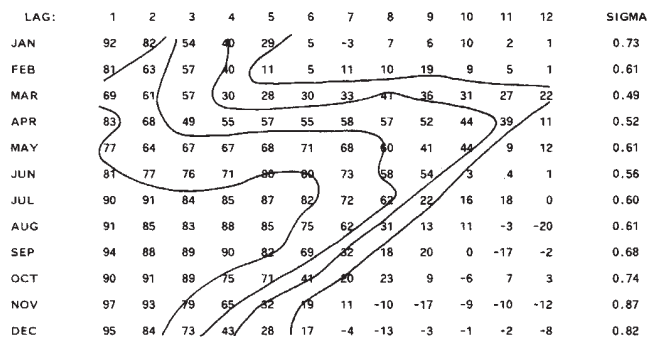


Fig. 2 Lag correlations ($\times 100$) of SST index with itself in subsequent months, 1947-69. For example, January with February is 0.92, August with December 0.88. Contours are for 0.8, 0.6, 0.4 and 0.2. Right-hand column, s.d. of index ($^{\circ}\text{C}$).

central sector (column 2) shows intermediate behaviour. Although the general levels of correlation are different in the three sectors, the annual cycles are similar, so that over the region as a whole (column 4) the correlation ranges from strongly positive in December to February to near zero in June to November. Because the SST index is highly negatively correlated with the strength of the Walker circulation, the values in column 4 are an estimate of C .

I hypothesize that the annual variation in the contribution of the cloudiness/radiation mechanism to p can account for the observed annual variations displayed in Fig. 2. To test this, I use the following version of the feedback model:

$$\Delta A = -0.17(A - B - R_a)$$

$$\Delta B = -0.033(B - pA - R_b)$$

$$p = p_0 + p_1 C$$

where R_a , R_b are Markov series for which the n th term is given by

$$R_{a,n} = 0.5R_{a,n-1} + \epsilon_{a,n}$$

$$R_{b,n} = 0.7R_{b,n-1} + \epsilon_{b,n}$$

and $\{\epsilon_a\}$, $\{\epsilon_b\}$ are series of random numbers normally distributed with standard deviations 0.1, 0.02 respectively. ΔA , ΔB represent the changes in A and B in a time step of 1 day. As a trial, let $p_0 = 1.02$, $p_1 = 0.02$, the values being chosen such that the mean and range of p are approximately the same as I used previously¹. The only seasonal variation in the model is that of the cloudiness regressions C . The model is run for 100 simulated years.

Figure 3a shows the pattern of persistence and variability of the model variable B , and should be compared with Fig. 2. Agreement in the general level of correlations and in the amplitude of the seasonal variation is expected because of the choice of values for p_0 and p_1 , and the question of modifying the constants to obtain the best possible match was not pursued. The phase information in Figs 2, 3a is, however, based on independent data. A comparison indicates a strong similarity

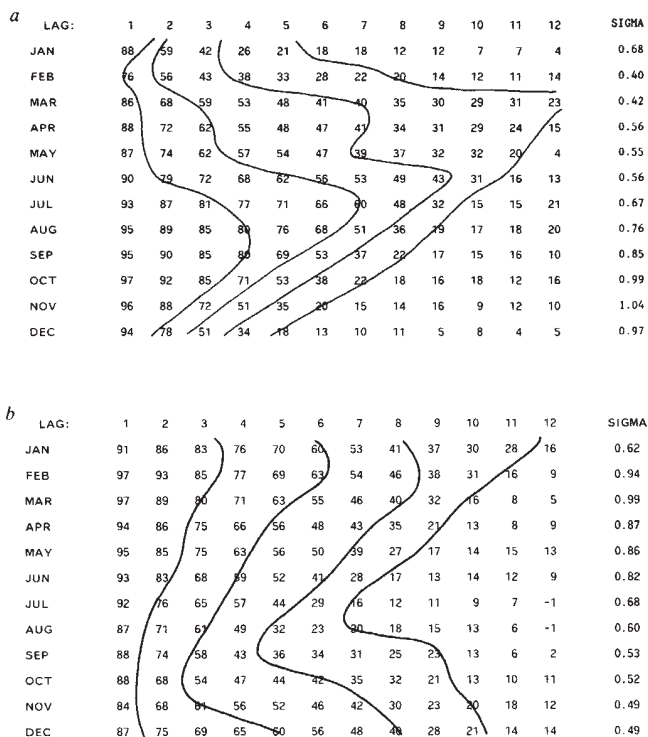


Fig. 3 a, Lag correlations ($\times 100$) of index B (output of the feedback model) with itself in subsequent months. Data for 100 yr. Right-hand column: s.d. of B (arbitrary units). b, As a but based on the hypothesis that increased cloudiness favours warmer sea.

of the seasonal patterns of both persistence and variability. As a measure of agreement, $C_3 = 0.84$, and the correlation between the sigma values (C_s) is 0.86. The only appreciable disagreement is around February-April, when the model phase is about a month earlier than the observed phase. To confirm the stability of the model results, two further runs were made using different random series, resulting in C_3 values of 0.76 and 0.91, and C_s values of 0.90 and 0.91.

These results are consistent with the hypothesis that the persistence in the SO is due to positive feedback, and that the cloudiness/radiation mechanism plays a substantial role in the seasonal modulation of the persistence by reducing the net positive feedback during the northern winter. Hastenrath and Wu⁵ have proposed that a similar association between SST and cloudiness in the Australasia region may be involved in a different aspect of the SO, the interannual fluctuations.

As a further demonstration, I test whether the hypothesis would hold if increased cloudiness favours warmer sea, by taking $p_0 = 0.66$, $p_1 = 0.02$, where p_0 was chosen such that the mean of the 12 values of p would be the same, and the positive sign for p_1 represents the new assumption. A 100-yr simulation generated Fig. 3b. Figure 3b is very different from Fig. 2 with respect to phase. Therefore, if increased cloudiness favours warmer sea, the hypothesis that the cloudiness plays a major role in the feedback is not supported.

Finally, we may ask why the relationship of cloudiness to SST does vary in the observed fashion. Only partial explanations can at present be suggested. The cloud in the west of the region is mainly convective and therefore would tend to increase in association with the increase of convergence that would accompany higher SST. In the east of the region, on the other hand, the normally cold sea is associated with stable air and stratocumulus cloud, especially during the season centred on September when the sea is coldest⁷. When the sea is warmer than normal in this area, increased instability will tend to break up the cloud deck. Perhaps, during the northern winter, the prevalent cloud type tends to become convective and thereby changes the sign of the relationship between anomalies.

A recent general circulation model experiment⁸ has shown reasonable success at simulating the cloudiness anomaly pattern associated with an SST anomaly in January, but similar experiments for other months have not yet been reported. The importance of correct representation of radiative processes for the simulation of atmospheric features has recently been demonstrated⁹. If cloudiness plays as important a role in air-sea interaction as the present results suggest, accurate simulation of it will be essential before good simulations of SST variations, and therefore the SO persistence, can be achieved.

Received 28 December 1983; accepted 9 May 1984.

1. Wright, P. B. *Nature* **277**, 371-374 (1979).
2. Wright, P. B. *Mon. Weath. Rev.* (in the press).
3. Egger, J., Meyers, G. & Wright, P. B. *Mon. Weath. Rev.* **109**, 1139-1149 (1981).
4. Liebmann, B. & Hartmann, D. L. *J. Atmos. Sci.* **39**, 1153-1162 (1982).
5. Hastenrath, S. & Wu, M.-C. *Arch. Met. Geophys. Bioklim. B* **31**, 1-37 (1982).
6. Wright, P. B. *Mon. Weath. Rev.* **107**, 1567-1580 (1979).
7. Horel, J. D. *Mon. Weath. Rev.* **110**, 1863-1878 (1982).
8. Blackmon, M. L., Geisler, J. E. & Pitcher, E. J. *J. Atmos. Sci.* **40**, 1410-1425 (1983).
9. Ramanathan, V., Pitcher, E. J., Malone, R. C. & Blackmon, M. L. *J. Atmos. Sci.* **40**, 605-630 (1983).

Modelling radioactive waste disposal by penetrator experiments in the abyssal Atlantic Ocean

T. J. Freeman*, C. N. Murray†, T. J. G. Francis‡, S. D. McPhail‡ & P. J. Schultheiss‡

* Building Research Establishment, Department of the Environment, Building Research Station, Garston, Watford WD2 7JR, UK

† Division of Chemistry, Commission of European Communities Joint Research Centre, 21020 Ispra, Varese, Italy

‡ Institute of Oceanographic Sciences, Wormley, Godalming, Surrey GU8 5UB, UK

The use of streamlined projectiles is a possible method of disposing of high-level radioactive waste in oceanic sediments¹⁻³. But there is little information about the manner in which freely-falling projectiles impact on the ocean floor^{4,5}. To correct this situation, the Engineering Studies Task Group of the OECD's Nuclear Energy Agency (NEA) Seabed Working Group, through which research into the feasibility of high-level radioactive waste disposal in the sediments of the ocean floor is coordinated, recommended a programme of studies with instrumented model penetrators⁶. The first tests in this programme are reported here and were carried out on RRS *Discovery* cruise 134 in the Great Meteor East (GME) radioactive waste disposal study area in March 1983, as a collaborative experiment between the Building Research Establishment, the Joint Research Centre and the Institute of Oceanographic Sciences.

The GME area lies in the southwestern distal part of the Madeira Abyssal Plain (Fig. 1). Extensive geological and geophysical studies have been carried out there by the Rijks Geologische Dienst of the Netherlands and by the UK Institute of Oceanographic Sciences⁷⁻¹¹. The sediments found there consist of fine-grained turbidites of calcareous ooze, typically a couple of metres thick, interbedded with thinner pelagic units ranging from ooze to pelagic clay¹¹. The sand component is negligible in the west of the area but sandy beds appear towards the east from about 24° 30' W (ref. 8).

Four similar penetrators were dropped in the GME area, each measuring 3.25 m in length and 0.325 m in diameter (Fig. 2). They were constructed of solid steel and weighed 1.8 tonnes in air. The design of the penetrators was based on a detailed hydrodynamic analysis of several possible configurations of nose shape, length/diameter ratio, and stabilizing surfaces carried out by Aeromacchi spa¹². The objective of the present study was to design a penetrator which would be stable during its passage through 5 km of water and would achieve maximum possible penetration on impact with the sediment. Penetrator drag, ter-

Table 1 Breakdown of drag coefficient of the penetrator in the water¹²

V (m s ⁻¹)	10	20	30	40	50
Friction body	0.107	0.096	0.091	0.088	0.085
Friction fins	0.032	0.027	0.026	0.025	0.024
Afterbody	0.025	0.025	0.025	0.025	0.025
Roughness	0.016	0.015	0.014	0.014	0.014
Total	0.180	0.163	0.156	0.152	0.148

minimal velocity, hydrodynamic loads and path stability were considered. Various conclusions for the penetrator design shown in Fig. 2 were reached.

(1) A terminal velocity of about 50 m s⁻¹ would be reached, corresponding to an overall drag coefficient of 0.148. The breakdown of drag components at different velocities is shown in Table 1.

(2) The distance necessary to reach terminal velocity would be about 600 m.

(3) The motion produced by any disturbance to the vehicle's angle of attack would be of a damped oscillatory type, with damping greater than critical. Such a disturbance would be reduced to one-tenth of its original value in a travel length of 10 m. The only likely cause for such a disturbance would be current shear. A step increase in horizontal current of 1 m s⁻¹ (2 knots), which is very unlikely, would create an instantaneous angle of attack of only 1° for a penetrator travelling at 50 m s⁻¹.

(4) Safety considerations dictated that the penetrators be launched horizontally. Due to the high stability characteristics of the penetrator, a rapid rotation from the horizontal to the vertical position was to be expected. This was confirmed by calculation of the trajectory of a horizontally-dropped penetrator. After 2 s the penetrator would be within 6° of the vertical and moving at 12 m s⁻¹. After 8 s it would be 2° from the vertical, moving at 42 m s⁻¹, at a depth of 190 m, and horizontally displaced from its release point by 13 m.

Before launch each penetrator was fitted with an acoustic transmitter, housed in a cylindrical recess in the tail, which emitted a constant 12-kHz signal as soon as it was immersed in the sea. As the penetrator falls the frequency f' of the signal received at a fixed hydrophone is given by:

$$f' = \frac{V_s}{V_s + V} f$$

where f is the frequency transmitted, V_s the velocity of sound in the medium surrounding the transmitter, and V the velocity of the source away from the hydrophone. The receiving hydrophone was suspended beneath a spar buoy floated away from the ship's side and, after amplification, conveyed to a shipborne receiving unit which linearly converted the Doppler frequency shift to d.c. voltage. This voltage output was recorded together with an accurate time code on magnetic tape and, for immediate observation during the experiment, by a jet-pen recorder. A record of Doppler-shifted frequency against time was thus obtained.

The results of the four penetrator drops are summarized in Table 2. The first two penetrators were dropped about a kilometre apart in a location where previous coring and the ship's 3.5-kHz profiler records indicated negligible quantities of sand in the upper part of the sedimentary column. Good penetration on 3.5-kHz records is obtained in regions with little or no coarse sediment¹³; that achieved at the first site was 80 m. Very similar records of frequency versus time were obtained from the first two impacts, with the acoustic signal being detected to the full penetration depth (at which the signal frequency returned to its shipboard value, indicating zero Doppler shift) in each case; it was thus decided to move to a second site 65 km to the east for the remaining two drops, where a sandy layer was known to be present in the sediment column from the Dutch coring⁸. The acoustic penetration achieved with the 3.5-kHz profiler at

Table 2 Summary of the four penetrator drops

Trial no.	Time 20 March 1983	Position		Water depth (m)	Descent time through water (s)	Impact velocity (m s ⁻¹)	Tail penetration (m)
		Lat. (°N)	Long. (°W)				
1	09.50	31° 25.5'	24° 44.2'	5,433	115.5	46.4	30.0
2	10.46	31° 25.6'	24° 44.7'	5,433	118.5	48.3	31.3
3	15.44	31° 23.6'	24° 02.6'	5,425	110	50.9	~27
4	16.22	31° 23.5'	24° 02.8'	5,425	115.5	47.3	Unknown

this location was 40 m. The penetrators were dropped but the signal from the last one was lost soon after impact and its penetration history was not recorded.

The terminal velocities of the penetrators through the water column were found to be in good agreement with the value predicted by the Aermacchi design study. Velocity–depth plots for three of the penetrators are shown in Fig. 3. Frequency was converted to velocity using the equation given above and the sound velocity profile for the water column in the vicinity¹⁴. Depths were calculated by integrating the penetrator velocities and working backwards from the time of impact. The velocities of all the penetrators were observed to decrease with depth once terminal velocity had been achieved. A decrease of ~4% is to be expected due to the increasing viscosity and density of the water with depth. An additional factor contributing to the velocity variation with depth might be deviation of the penetrator from a vertical path due to small misalignments of the fins. However, the acoustic tracking system was not capable of distinguishing such deviations, observing only velocity away from the hydrophone.

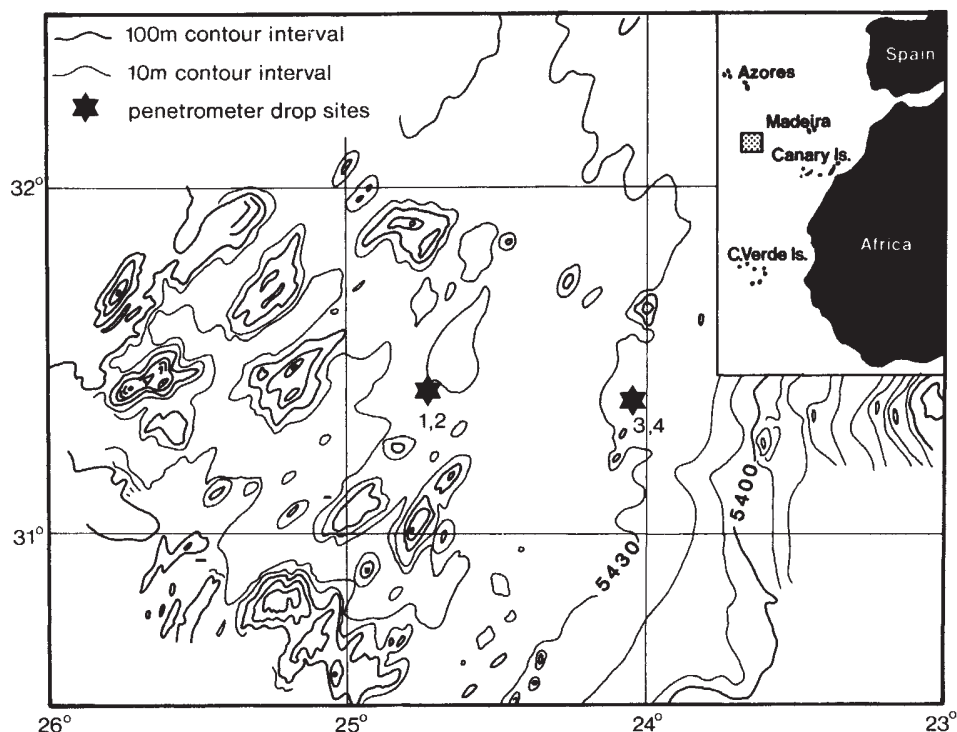
The velocity history of one of the penetrator impacts is shown in Fig. 4a. To obtain this record the sound velocity in the sediment was assumed to be 1,552 m s⁻¹, equal to that in the bottom water¹⁴. Measurements of sound velocity made at room temperature and pressure on board the ship, on cores from the vicinity of the drop, found sound velocities in the surface sediments 25 m s⁻¹ less than in seawater. The lower sound velocity in the sediment probably accounts for the apparent acceleration of the penetrator after hitting the sediment. What is observed is the increased Doppler shift due to the lower velocity of the

medium. If this is so, the positive 'blip' on the record corresponds to the moment at which the sediments closed over the tail of the penetrator. This may have occurred when the tail was level with the sea floor or when it was at some depth into the sediment. The point of impact marked in Fig. 4a assumes that the increase in Doppler shift occurred when the penetrator's tail was level with the sediment surface. This Doppler shift could be interpreted as evidence for hole closure behind the penetrator. However, the physics of a sound source travelling at speed down a borehole are complex and other interpretations cannot be ruled out.

The spikes shown on Fig. 4a do not correspond to real velocity changes but are the result of temporary losses of signal during the penetration. Similarly the greater noise level observed when the penetrator is stationary compared with the period immediately before impact is due to the attenuation caused by burial beneath 31.3 m of sediment. One of the surprises of the first two drops was the remarkably low attenuation of the acoustic signal by the sediment. For the first penetrator the attenuation was only 3 dB, corresponding to 0.1 dB m⁻¹ through the sediment; for the second penetrator it was 13 dB, corresponding to 0.4 dB m⁻¹. These values are lower than reported attenuation rates for 12 kHz sound in comparable fine-grained marine sediments¹⁵.

Before interpreting the velocity–time curve for each penetrator, the records were processed to remove spikes, being smoothed and corrected for the different sound velocity in the sediment, which was assumed to be 25 m s⁻¹ lower than in the bottom water. Integration of the velocity–time curve from the moment of impact to the time at which zero velocity is reached gives the tail penetration depth of the penetrator (Table 2).

Fig. 1 Bathymetric map of the GME radioactive waste disposal study area showing the positions of the penetrator drops.



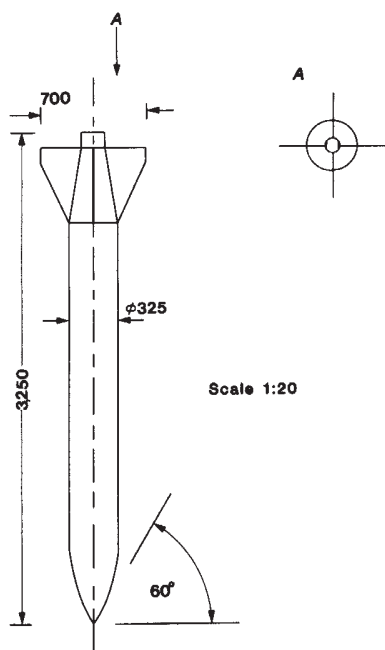


Fig. 2 The model penetrator used in the experiments.

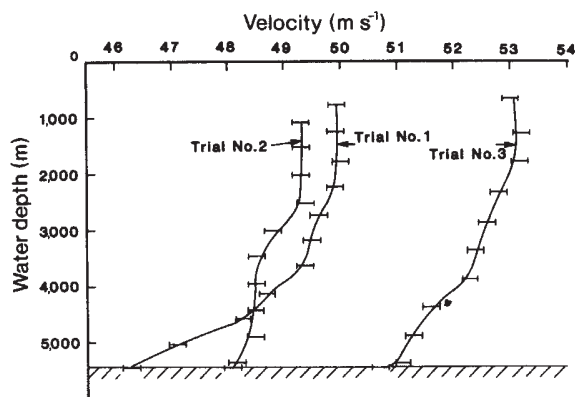


Fig. 3 Terminal velocities of the first three penetrators through the water column.

Full interpretation of the penetrator data requires an, as yet unavailable, detailed knowledge of the geotechnical properties of the sediments. However, shear strength measurements have been reported for a 9-m long piston core recovered 75 km from the first penetrator site⁸. These agree broadly with a shear strength profile of $C_u = 1.5z$, where C_u is the undrained shear strength in kPa and z is the depth below sea floor in metres, which has been suggested for oceanic sediments¹⁶, although considerable variation from a linear relationship does exist. In particular the pelagic deposits were appreciably stronger than the turbidite layers.

Even with a knowledge of the geotechnical properties, methods for calculating the penetration of projectiles into sediments are still being studied, those used so far being largely empirical¹⁶⁻²¹. We adopted a very simple approach based on studies carried out by Arup¹⁶ in which resistance to the penetration is a function solely of the undrained shear strength of the sediment:

$$\text{Base resistance} = N_{CD} C_u A_f$$

$$\text{Shaft resistance} = \alpha_d C_u A_s$$

where A_f and A_s are frontal and surface areas of the penetrator respectively, N_{CD} is an empirical dynamic end bearing capacity factor and α_d is an empirical dynamic shaft adhesion factor. Equating the sediment resistance plus buoyant forces to the rate

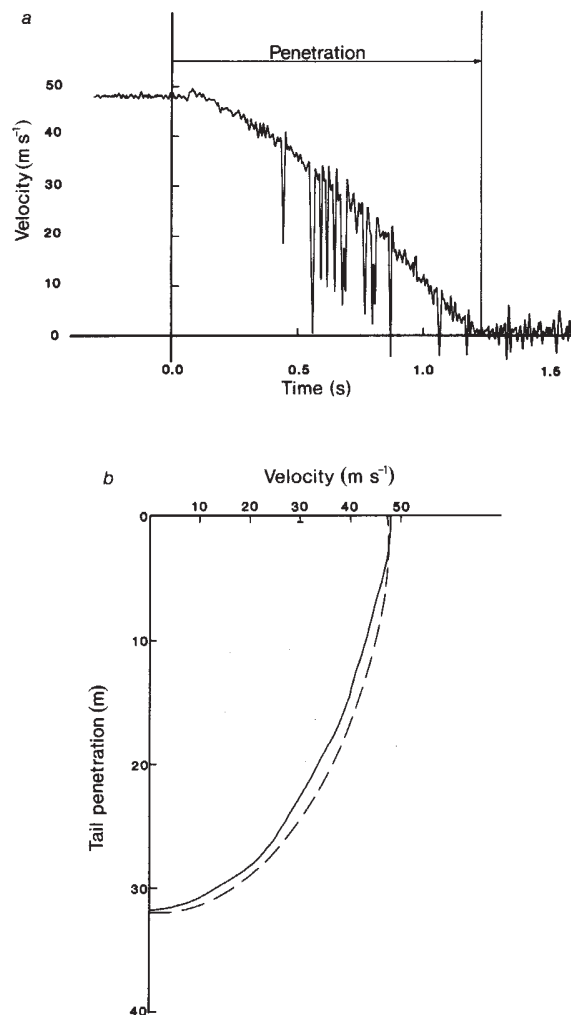


Fig. 4 a, Penetrator velocity plotted against time for the second penetrator. Velocity was obtained from the Doppler-shifted frequency using a sound velocity in the sediment of $1,552 \text{ m s}^{-1}$. b, Smoothed velocity-depth profile for the second penetrator derived from a. The predicted profile (dashed line) was calculated assuming a linear increase of shear strength with depth.

of change of momentum of the penetrator, the deceleration history was calculated by a computer program. With $N_{CD} = 20$, $\alpha_d = 0.4$ and the shear strength profile $C_u = 1.5z$, good agreement was obtained with the observed deceleration curve (Fig. 4b). Although this analysis has obvious shortcomings, in particular that soil resistance is taken to be independent of the penetrator velocity, the shape of the deceleration curve is predicted remarkably well.

Further model penetrator drops are planned at several oceanic locations in the next few years. Future experiments will determine the influence of sediment type and penetrator shape on the penetration achieved. A 3.5-kHz transponder is under development which, it is hoped, will provide a communication link between buried penetrators and the surface ship. This will allow a range of measurements of the environment around the penetrator to be made, culminating in studies of what fills the disturbed zone created by the entry into the sediment. A fuller account of the experiment reported here will be published elsewhere.

This research was carried out under contract to the Commission of the European Communities and to the UK Department of the Environment, as part of their radioactive waste management research programmes. The results will be used in the formulation of UK government policy, but at this stage they do not necessarily represent government policy.

Received 13 February; accepted 23 May 1984.

- Valent, P. J., Lee, H. J., True, D. G. & Malloy, R. J. in *High-level Radioactive Waste Management Alternatives* Vol. 3 (eds Schneider, K. J. & Platt, A. M. 6B1-6B42) (Battelle, Pacific Northwest Laboratories, 1973).
- Valent, P. J. & Lee, H. J. *Mar. Geotech.* **1**, 267-293 (1976).
- Silva, A. J. *Oceanus* **20**, 31-40 (1977).
- Whitmarsh, R. B. & Lilwall, R. C. *Proc. Oceanol. Int. Pap.* 4.2 (Spearhead Publications, Kingston-upon-Thames, 1982).
- Beard, R. M. *Expendable Doppler Penetrometer: a Performance Evaluation* (US Naval Civil Engineering Laboratory, Port Hueneme, Rep. TR 855, 1977).
- Talbert, D. M. *Rep. 3rd Interim Meeting of the Seabed Working Group ESTG* (Sandia Labs Rep. no. SAND 82-2701, 1983).
- Kuijpers, A. *The Sedimentology of the Two Northeast Atlantic Study Areas: the Western Madeira Abyssal Plain and the Area West of Great Meteor Seamount* (Rijks Geologische Dienst Prog. Rep., 1981).
- Duin, E. J. Th. & Kuijpers, A. *Geological Studies on Abyssal Plains in the North Atlantic* (Rijks Geologische Dienst Prog. Rep., 1982).
- Francis, T. J. G. *et al. Geophysics and Sediment Sampling in the Northeast Atlantic* (Institute of Oceanographic Sciences Cruise Rep. 117, 1981).
- Francis, T. J. G. *et al. Geological and Geophysical Studies between the Azores and the Canaries* (Institute of Oceanographic Sciences Cruise Rep. 141, 1983).
- Weaver, P. P. E. & Kuijpers, A. *Nature* **306**, 360-363 (1983).
- Aermacchi spa: Hydrodynamic Analysis and Design of High-level Radioactive Waste Disposal Model Penetrator* (Rep. 419/STE/82, Commission of European Communities Joint Research Centre, Ispra, 1982).
- Damuth, J. E. *Mar. Geol.* **38**, 51-75 (1980).
- Saunders, P. M. *CTD Data from the North Madeira Basin, 19 January-12 February 1981, Discovery Cruise 117* (Institute of Oceanographic Sciences Data Rep. No. 26, 1981).
- Hamilton, E. L. *J. acoust. Soc. Am.* **59**, 528-535 (1976).
- Arup, O. *et al. Penetrator Option-studies Relevant to Emplacement in Deep Ocean Sediments* (DOE Rep. no. DOE/RW/82.102, 1982).
- Young, C. W. *Development of Empirical Equations for Predicting Depth of Earth Penetrating Projectiles* (Sandia Labs Rep. no. SC-DR-67-70, 1967).
- McNeill, R. L. *Oceans 81 Conf.*, Boston, 688-693 (IEEE, New York, 1981).
- Burchett, S. M. *Subseabed Disposal Program, A. Rep.* (Sandia Labs Rep. no. SAND 82-0664/1, II, 2, Appendix HH, 359-392, 1982).
- Dzwilewski, P. T. *Subseabed Disposal Program, A. Rep.* (Sandia Labs Rep. no. SAND 82-0664/1, II, 2, Appendix GG, 335-358, 1982).
- Canfield, T. & Burchett, S. N. *Subseabed Disposal Program, A. Rep.* (Sandia Labs Rep. no. SAND 82-0664/1, II, 2, Appendix FF, 301-334, 1982).

Discovery of a large brine deep in the northern Red Sea

Guy Pautot*, Pol Guennoc†, Alain Coutelle* & Nikos Lyberis§

* CNEXO and † BRGM, Centre Océanologique de Bretagne, BP 337, 29273 Brest cédex, France

‡ G. I. S. 'Océanologie et Géodynamique', Université de Bretagne Occidentale, 29283 Brest, France

§ Laboratoire de Géodynamique, 4 place Jussieu, Tour 15, 75230 Paris cédex 05, France

The Red Sea is formed of marginal areas, shallower than 1,000 m, which are separated by a deeper axial zone. The axial zone can be subdivided into three main segments. The southern segment is a relatively simple trough which, north of 16° N, corresponds to an oceanic rift¹. The middle segment, from 19°30' N to 23°50' N, is characterized by the presence of about 15 closed depressions or axial deeps, ranging from several kilometres to less than a kilometre in length, and containing hot or cold brines underlain by their associated deposits of soft- and fine-grained metalliferous muds¹⁻³. The northern segment of the axial zone is shallower, with depths rarely exceeding 1,500 m, and has a smoother relief than the southern and central segments. In the northern segments, south of the Al Wajh depression at 25°25' N, the axis strikes north-south (Fig. 1). North of the depression, towards the approaches of the Gulf of Suez, the strike is NNW-SSE. The small, circular (1 km diameter) Oceanographer deep is located⁴ along the western border of the axial zone of this northernmost segment. The Jean Charcot deep, which we discovered on 3 December 1983 during a cruise of the RV *Jean Charcot*, lies near 25°15' N, 35°22' E, close to the axis of maximum depth of the same segment. In contrast with the magnetically quiet neighbourhood, a strong dipolar magnetic anomaly typifies this deep, which in spite of intensive international surveys, is the first brine deep to have been discovered in the Red Sea in the past 10 years.

The presence of brine accumulation was recognized by its characteristic signature on standard 3.5- and 12-kHz echograms. A Seabeam (multi-beam echosounder), magnetic, gravity and

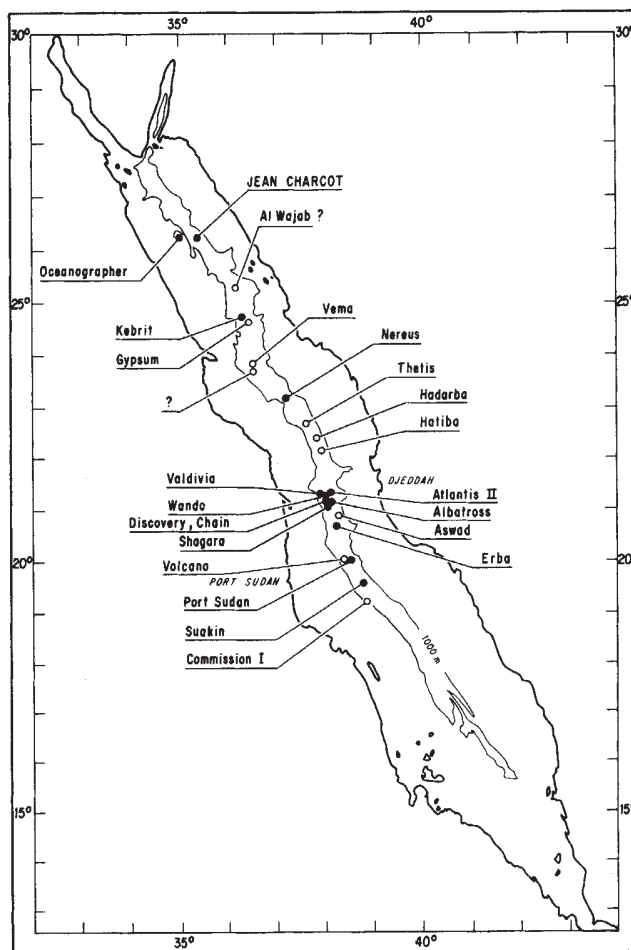


Fig. 1 Schematic map of the Red Sea area showing the axial depression (as defined by 1-km isobaths) and the main axial deeps or depressions containing hot brines and metalliferous sediments (●), or containing only metalliferous sediments (○). The Jean Charcot deep is also indicated. The 1-km isobath outlines the main segments (oriented from N-S to NNW-SSE) of the Red Sea axial zone.

single-channel reflection seismic survey was then conducted. The Jean Charcot deep is approximately rhombic in shape and is 10 km long and 6 km wide (Fig. 2). The long axis of the deep strikes 138° and is represented by a central topographic high, which is also the direction of the alignment of several small, presumably volcanic, cones located on the crest of the high, one of which marks its highest point (-900 m). The deep comprises four basins separated by the main axial ridge which is flanked by two lateral saddles. The greatest depth (1,490 m, uncorrected value assuming a velocity of 1,500 m s⁻¹) is in the southern basin. The slopes of the central relief and of the walls of the depression are generally 40% or more. Outside the borders of the deep, the seafloor is topographically smooth and is shallower than in the deep itself: about 1,200 m and 1,100 m to the SE and NW respectively. On the ESE side is a linear ridge and valley structure, oriented roughly NNE-SSW, which is a transform fault direction.

Brines as thick as 250 ms (approximately) double travel time are clearly seen in the two larger southern basins (Fig. 3) where their upper surfaces are marked by a strong reflector near 1,300 m (uncorrected value). Similar flat echoes occur at the same depth (very near the sea floor) in the two smaller and shallower northern basins, indicating the presence here of very thin brines. A water sample was taken of the brines 25 m above the sea floor of the southern basin. The potential temperature measured was 24.8°C, that is, 2 or 3°C higher than that of normal bottom water. This is one of the lowest temperatures recorded for the

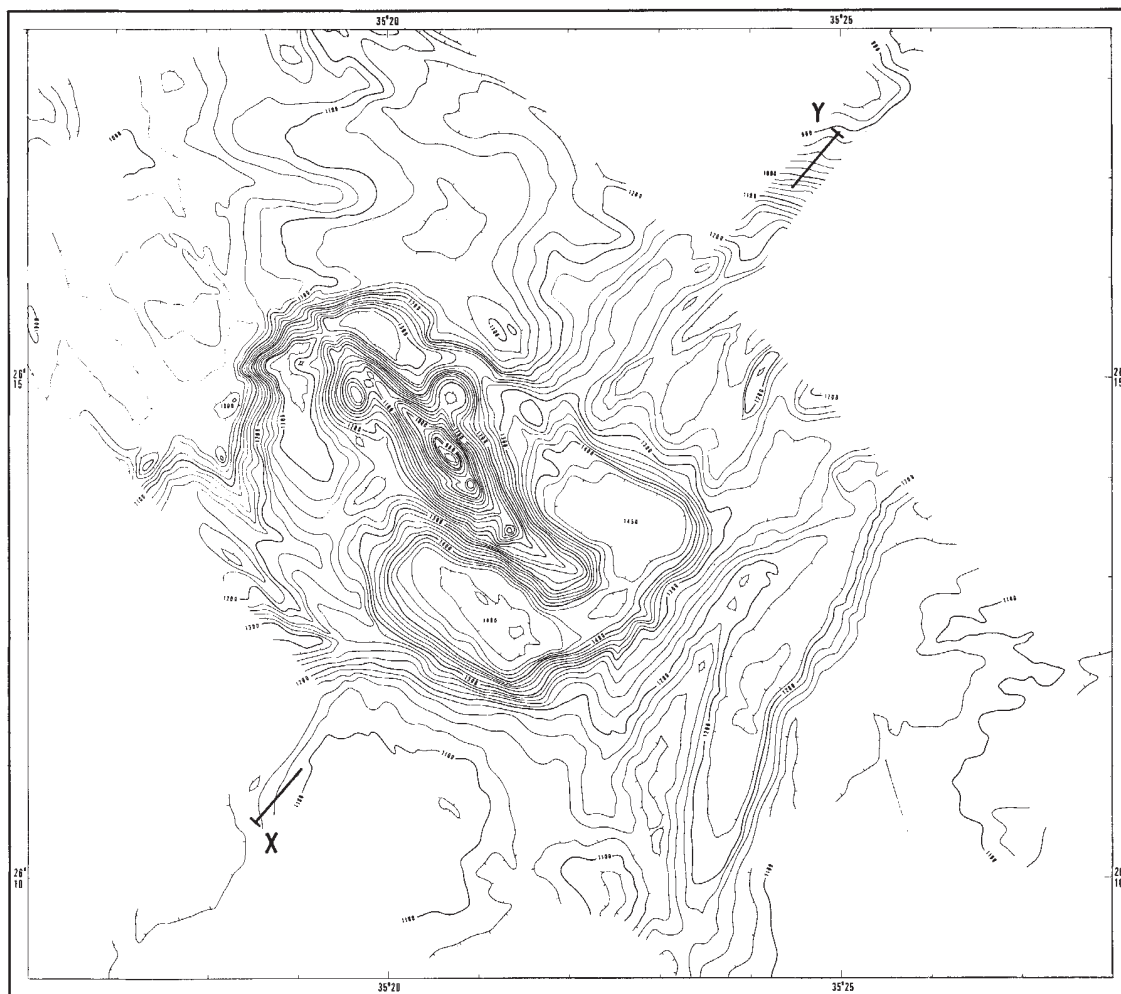


Fig. 2 Seabeam bathymetric map of the Jean Charcot deep. This parallelepiped-shaped depression is orientated NNW-SSE, which corresponds to the median direction of the Red Sea northern axial segment. The dimensions of the deep are about 10×6 km. An axial crest with small seamounts, and two lateral highs over which high-amplitude magnetic anomalies were recorded, limit the four basins. The maximum depth (1,490 m, uncorrected) has been observed in the south-west basin. The northern basins are more restricted and shallower: maximum depths are $\sim 1,310$ m. Brines are observed in the two southern basins (see Fig. 3), the top surface of which must lie at about 1,310 m. The thin presence of brines is suspected on the north-west basin. Note the well-defined transverse (transform) morphology (valley and ridge) on the southern border. X-Y refers to the seismic line section shown on Fig. 3.

Red Sea brines, which have temperatures varying between 23 and 62.3 °C (maximum temperature in the Atlantis II deep lower brines^{3,5,6}).

The measured salinity of the brines we sampled is 310‰ and the chlorinity is 190‰. These values are similar to those recorded in the lower brines of the Atlantis II deep.

A gravity core, 800 cm long, was taken in the southern basin. The core comprises four hydrothermal-type sedimentary units separated by three dominantly biotrital units. The hydrothermal material shows a centimetric, regular bedding of brown and green layers. The biotrital sequences were formed by beige-coloured oozes. The upper biotrital unit presents gravity slides and at the base, a decimetric bed of volcanoclastic material. X-ray diffraction analyses of the sediments show that they are mainly composed of calcite + dolomite + quartz with traces of kaolinite + plagioclase + illite + pyrite and some amorphous material. An ICP analysis reveals (Table 1) a high SiO_2 content, which is related to the presence of abundant radiolaria, and a Zn enrichment (0.3–0.6%) similar to the mean Zn content of the most mineralized deeps^{3,7,8}. The concentrations of the other elements are quite similar to those of other Red Sea biotrital sediments⁹ and of the basal detrital unit of the Atlantis II deep^{8,10}.

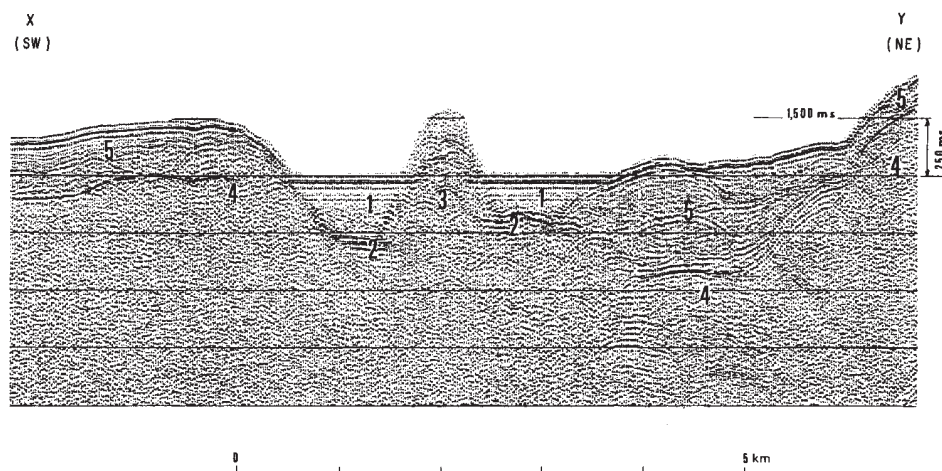
The sediment collected directly overlies the volcanic basement and a fragment of basalt was found in the core catcher that had

been damaged during the coring. A preliminary observation of the rock fragments shows that it is a hyalobasalt. The glass is slightly oversaturated in Si (5% normative quartz) and rich in TiO_2 (3%) and Fe (FeO^* 3%), the water content being $\sim 2\%$. The microlites are composed of plagioclase, whose composition is at the limit between andesine and labrador, and of Mg-olivine and clinopyroxene (salite). The dominant phenocrysts are plagioclases of the same composition as the microlites. Some metastable relicts of xenomorphic amphiboles are also present. The chemical analysis of the salites shows little variation of composition. Using the discriminative diagram of Le Terrier *et al.*¹¹, an alkaline or transitional affinity can be attributed to these rock fragments.

The basalt sample taken from the bottom of the deep, the small cones aligned on the central relief (Fig. 2), the strong magnetic anomaly observed within this area (about 500 nT from peak to peak) and the lack of sediment over it, lead us to conclude that the central ridge within the Jean Charcot deep is a young volcanic edifice built during a recent opening in the northern Red Sea. Its location in the centre of the axial area and its alignment along the regional direction (about 140° N) of the northern axial segment (Figs 1 and 2) of the Red Sea strengthens this hypothesis.

The magnetic anomaly map of the area of the Jean Charcot deep (in ref. 12) shows no linear trends but a broad dipolar,

Fig. 3 SW-NE seismic profile across the Jean Charcot deep (for localization, see Fig. 2). Source, water gun; speed, 10 knots. The vertical exaggeration is about 6 (for seawater velocity $1,500 \text{ m s}^{-1}$). 1, Brines. The top of the brines is shown as a characteristic flat echo. Stratification is seen well inside the brines. 2, Strong echoes representing recent metalliferous sediments. 3, Crest of central ridge devoid of any sediments and more probably of volcanic origin. 4, Upper Miocene evaporites. 5, Plio-quaternary biotrital sediments. In the Red Sea, the two units are classically separated by a strong reflector, or group of reflectors, called reflector S. On the north-east border of the deep the layered series between two strong reflectors (between 5 and 4) may correspond either to a thicker Plio-Quaternary unit (basin filling) or to Upper Miocene stratified series (evaporitic marls?).



axial anomaly which is the most prominent in the northern Red Sea. These dipolar anomalies have been interpreted, together with other geophysical evidence, as being due to localized bodies intruded into a thin crust, the nature of which (thinned and intruded continental crust¹³⁻¹⁵ or oceanic crust¹⁶) is controversial.

Table 1 Inductive coupled plasma analysis of two samples of deepest sediments (dry salt-free) in a 800-cm long gravity core taken in the south-west basin of Jean Charcot deep

	1	2
(%)		
SiO ₂	56.9	52.6
Al ₂ O ₃	6.7	8.9
Fe ₂ O ₃	2.7	3.7
CaO	14.0	13.1
MgO	2.5	2.4
K ₂ O	1.1	—
MnO	0.04	0.05
TiO ₂	0.53	0.82
(p.p.m.)		
P ₂ O ₅	766	1,111
Li	—	—
Be	—	—
B	34	11
V	63	92
Cr	57	60
Co	97	53
Ni	426	230
Cu	44	39
Zn	2,959	6,065
As	108	111
Sr	384	302
Y	—	—
Zr	108	103
Nb	—	—
Mo	20	14
Ag	—	—
Cd	23	20
Sn	—	—
Sb	—	11
Ba	261	197
La	—	—
Ce	25	20
W	—	—
Pb	49	47
Bi	—	—
Total + loss on fire	84.5%	78.2

Analyses from Bureau de Recherches Géologiques et Minières, Orléans, France.

Most authors agree with the conclusion that the northern Red Sea is at an earlier stage of opening than the central and southern parts, where oceanic-type magnetic lineations have been clearly identified^{13,17,18}, and where typical oceanic tholeiitic basalts (mid-oceanic ridge basalts) have been recovered¹⁹⁻²¹. It is therefore interesting that the basalt sample we dredged provides the first evidence that the northern Red Sea basement has alkaline-transitional affinities, and it could well indicate the site at which the ocean originated.

In the general scheme that describes the propagation of a recent oceanic rift through deformed continental areas²², neo-oceanic crust is first created at distension points (nucleations) which are isolated from the tip of the rift. The localization of these points can be influenced by pre-existing lines of weakness such as marginal fracture zones which are present on the southern flank of the Jean Charcot deep.

The dimensions of the Jean Charcot deep are similar to those of other axial deeps such as the Nereus or Atlantis II deeps^{1,4}. This would indicate that the rate and the style of axial distension near 26°15' N, 23°10' N or 21°20' N are similar. This cannot be explained by a simple fan-shaped propagating oceanic opening model with its apex located between 23° and 24° N. If the rigidity concept of plate tectonics is applied, we must then consider either that the pole of rotation is located farther north than the position deduced from the fan-shaped oceanic opening model or that strike-slip movement occurred along the Arabia-Sinai plate boundary at the same time as the Jean Charcot deep was formed. Alternatively, the explanation could lie in non-rigid plate behaviour involving locally-varying opening rates within relatively short periods.

We thank Captain Y. Keranfec'h and the crew of the RV *Jean Charcot* for assistance; Y. Thisse for mineralogical and geochemical analyses; P. Thonon, R. Hékinian and R. C. Maury for petrological observations; and H. D. Needham for useful discussions. The cruise was supported by CNEXO. Contribution no. 866 of the Centre Océanologique de Bretagne.

Received 17 February; accepted 11 May 1984.

1. Bäcker, H., Lange, K. & Richter, H. *Geol. Jber D* **13**, 79-123 (1975).
2. Degens, E. T. & Ross, D. A. (eds) *Hot Brines and Recent Heavy Metal Deposits in the Red Sea* (Springer, Berlin, 1969).
3. Bäcker, H. & Schoell, M. *Nature phys. Sci.* **240**, 153-158 (1972).
4. Pautot, G. *Oceanol. Acta* **6**, 235-244 (1983).
5. Hartman, M. *Deep-Sea Res.* **27**, 161-171 (1980).
6. Monin, A. S. & Plakhin, E. A. *Deep-Sea Res.* **29**, 1271-1291 (1982).
7. Fletcher, A. W. & Mustafa, Z. in *National and International Management of Mineral Resources*, 3-16 (Institute of Mining and Metallurgy, London, 1980).
8. Thisse, Y. thesis, Univ. Orléans (1982).
9. Whitmarsh, R. B. *et al. Init. Rep. DSDP leg 23*, 1115-1123 (1974).
10. Bäcker, H. & Richter, H. *Geol. Rdsch.* **62.3**, 697-741 (1973).
11. Le Terrier, J., Maury, R. C., Thonon, P., Girard, D. & Marchal, M. *Earth planet. Sci. Lett.* **59**, 139-154 (1982).
12. Hall, S. A., Andreasen, G. E. & Girdler, R. W. *Direct. Gen. Miner. Resour. Bull.* **22**, F 1-F 15 (1977).

13. Izzeldin, Y. A. thesis, Univ. Strasbourg (1982).
14. Gerdes, K. D. & Styles, P. *Abstr. EGS-ESC Meet.*, Strasbourg, 81 (1982).
15. Cochran, J. R. *Am. Ass. petrol. Geol. Bull.* 67, 41-69 (1983).
16. Hall, S. A. thesis, Univ. Newcastle-upon-Tyne (1976).
17. Vine, F. J. *Science* 154, 1405-1415 (1966).
18. Roesser, H. A. *Geol. Jber.* D13, 131-156 (1975).
19. Chase, R. L. in *Hot Brines and Recent Heavy Metal Deposits in the Red Sea* (eds Degens, E. T. & Ross, D. A.) 122-128 (Springer, Berlin, 1969).
20. Coleman, R. G., Fleck, R. J., Hedge, C. E. & Eiment, E. D. *Direct. Gen. Miner. Resour. Bull.* 22, D 1-D 30 (1977).
21. Juteau, T. *et al. Bull. Centr. Rech. Explor. Prod. Elf-Aquitaine* 7.1, 217-231 (1983).
22. Courtillot, V. *Tectonics* 1, 239-250 (1982).

Zones of coastal hypoxia revealed by satellite scanning have implications for strategic fishing

Thomas D. Leming* & Warren E. Stuntz†

National Marine Fisheries Service, Southeast Fisheries Center, Mississippi Laboratories at * National Space Technology Laboratories, Mississippi 39529, USA, and † Pascagoula, Mississippi 39567, USA

Hypoxic (oxygen concentrations $<2.5 \text{ mg l}^{-1}$) and occasionally anoxic bottom waters occur along the inner continental shelf of Texas and Louisiana¹⁻⁴. Little is known, however, about the spatial and temporal scales of these hypoxic areas. Boesch⁵ concludes the nearshore continental shelf west of the Mississippi River Delta has characteristics potentially favourable for large-scale bottom water oxygen depletion. These characteristics include influences from river and marsh discharges of freshwaters, rich in nutrients and sediments. The river effluent flows over the higher salinity shelf waters, creating two-layer vertical density stratification. When surface waters are heated during the summer, vertical stability increases, especially during calm wind periods, and vertical eddy diffusion is limited. Large-scale phytoplankton blooms stimulated by nutrients from the freshwater runoff can then produce large amounts of organic matter which sink and, through respiration and decomposition, deplete oxygen concentrations in the bottom layer. The hypoxia is normally limited to several metres of the water column next to the bottom⁶ in water depths usually $<50 \text{ m}$. Because hypoxia appears to be related to surface chlorophyll and temperature, which can both be measured with the Coastal Zone Color Scanner (CZCS) on board the Nimbus 7 spacecraft^{7,8}, we have tried to determine if conditions favourable for formation of hypoxia could be detected and monitored from space. A linear discriminant function successfully identified areas of bottom water hypoxia detected by research vessels up to 10 days after satellite overpass. The discriminant function also successfully predicted hypoxic areas in June 1983 without resort to research vessel data. Shrimp and finfish were absent in the hypoxic zones; hence, the mapping from space of conditions favourable for hypoxia development may have significant marine resource implications for both strategic fishing and management.

A hypoxic bottom water condition off the south-west Louisiana coast in late June 1982 has been reported¹ from a cooperative Southeast Area Monitoring and Assessment Program (SEAMAP) cruise involving federal and state research vessels. The vessels occupied a series of randomly selected stations from Alabama to Texas. Each station included measurements at surface, mid-depth and bottom, of temperature, salinity, chlorophyll and dissolved oxygen, and timed tows of a 12.2 m headrope-length shrimp trawl. The vessels were off Louisiana on 14-25 June 1982.

On 14 June 1982, during orbit 18,374, Nimbus 7 passed over the northern Gulf of Mexico with the Louisiana hypoxic area in view of the CZCS. The CZCS data were first corrected for atmospheric Rayleigh scattering. Aerosol scattering was then estimated using the technique described by Gordon and Clark⁹; however, because this technique fails with high chlorophyll and sediment concentrations, the aerosol scattering estimate was

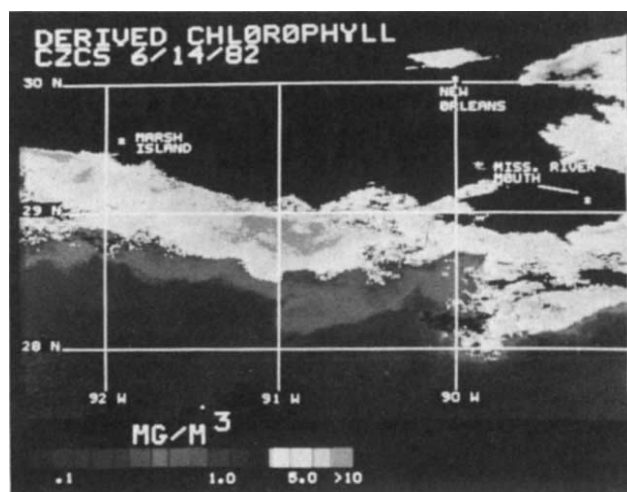


Fig. 1 CZCS-derived chlorophyll concentration for Nimbus 7 orbit 18,374, 14 June 1982. Land, clouds and heavily sedimented water are black.

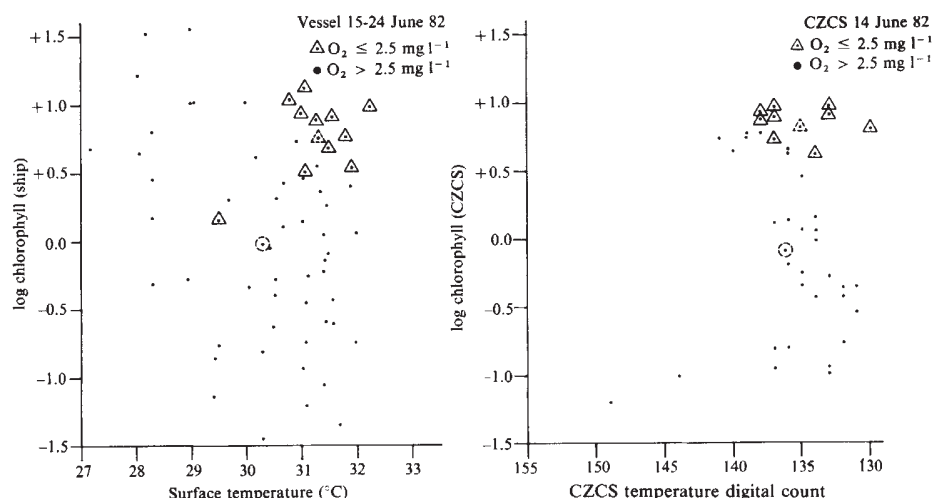
further corrected by iteration¹⁰. Chlorophyll concentrations were estimated based on corrected radiances at 443, 520 and 550 nm with a bio-optical algorithm¹¹⁻¹³. High sediment concentrations were excluded from processing by a conservative threshold in the 750 nm CZCS spectral channel. These areas were predominantly those affected by the Mississippi River plume and would include dissolved organic matter (yellow substance) of terrigenous origin. In areas of high chlorophyll concentration ($5-10 \text{ mg m}^{-3}$), effects of yellow substance on the chlorophyll algorithm are not serious¹⁴. Additionally, measurements throughout the Gulf of Mexico, including the Mississippi Delta area, have shown a strong covariance between the concentration of total suspended particulate matter and pigment concentration¹⁵.

A full resolution, derived chlorophyll image, resampled to a rectangular grid for the Louisiana coastal area on 14 June 1982, is shown in Fig. 1. Land, clouds and heavily sedimented waters are shown as black, and chlorophyll concentrations are shown ranging from dark grey ($<0.1 \text{ mg m}^{-3}$) to white ($\geq 10 \text{ mg m}^{-3}$). Derived measurements of chlorophyll concentrations exceeding 5 mg m^{-3} are evident in significant portions of the coastal waters extending from east of the Mississippi River Delta along the Louisiana coast. Unfortunately, no surface measurements of chlorophyll in cloud-free areas were available from the SEAMAP stations within 24 h of CZCS coverage for direct comparisons. However, later surface measurements (15-16 June) from areas where high chlorophyll concentrations were indicated ranged from 1.0 to 30.0 mg m^{-3} .

Vessel and satellite measurements suggested that waters in the region of low bottom oxygen concentrations were strongly stratified and experiencing a significant algal bloom. Additionally, these areas generally had sea-surface temperatures of 31°C or more. A plot of vessel measurements of surface chlorophyll and temperature (Fig. 2a) indicates that hypoxic and non-hypoxic stations were statistically separable. The F -ratio was 14.34 with 2 and 59 degrees of freedom, and the null hypothesis of equal vector group means can be rejected at $P < 0.001$. A similar plot for the subset of vessel stations in cloud-free areas of the 14 June CZCS image (Fig. 2b), using CZCS-derived chlorophyll and temperature digital counts, shows that the two groups are distinct. Digital count values were used for temperature because the calibration source onboard the CZCS is unstable, precluding absolute temperature calculations. For the CZCS data, $F = 10.75$ with 2 and 42 degrees of freedom and the null hypothesis is rejected at $P < 0.001$.

The separability of hypoxic and non-hypoxic stations based on CZCS-derived chlorophyll and temperature was used to build a linear discriminant function classifier using the data in Fig.

Fig. 2 *a*, Vessel-measured sea-surface chlorophyll concentration plotted against sea-surface temperature. Bivariate group means and standard deviations of log chlorophyll and temperature for hypoxic stations (dashed triangle) were 0.766 ± 0.285 and 31.30 ± 0.73 °C, respectively. For non-hypoxic stations (dashed circle) they were -0.021 ± 0.748 and 30.30 ± 1.25 °C. *b*, CZCS-measured chlorophyll concentration plotted against temperature digital count. Bivariate group means and standard deviations of log chlorophyll and temperature digital counts for hypoxic stations were 0.834 ± 0.103 and 135.8 ± 2.8 , respectively. For non-hypoxic stations, they were -0.094 ± 0.635 and 136.2 ± 4.2 .



2b. The CZCS digital temperature values were converted to standardized variables with mean of zero and standard deviation of one, allowing the coefficients to be used as predictors for other image dates. The coefficients of the linear discriminant function indicated that over 95% of the discrimination was accounted for by the derived chlorophyll. Vessel data have shown that hypoxia occurs in water depths ≤ 45 m. The ratio to total area of water areas with depth less than and greater than 45 m was used to adjust the discriminant function before classification for *a priori* probabilities that a given location would be hypoxic. The linear discriminant function was then used to classify the 14 June image into either hypoxic or non-hypoxic classes.

There was a good spatial relationship between SEAMAP stations reporting hypoxia and the region classified as hypoxic from CZCS data (Fig. 3a). The hypoxic SEAMAP stations are superimposed as white crosses on the classified image and the non-hypoxic stations as white dots. Of the 15 stations reporting hypoxia, 11 fell into or adjacent to the classified hypoxic area. Four stations were in an area of either cloud cover or heavily sedimented water. A few of the non-hypoxic stations fell on the edge or slightly into the classified hypoxic zone, but most (95%) classified correctly. Overall, the classification seemed to provide a good representation of hypoxic conditions off the coast of Louisiana and suggests that the areal extent was much larger than inferred from the vessels. The total hypoxic area shown in Fig. 3a is $\sim 6,000$ km². A hypoxic area contoured from vessel oxygen measurements alone was 1,200 km².

In 1983, CZCS digital tapes were acquired and processed within 24 h of satellite overpass during the SEAMAP survey. A relatively cloud-free image of the south-west Louisiana area was acquired on 9 June 1983. The chlorophyll image was produced and the coefficients of the discriminant function derived for the 14 June 1982 image were used to classify potentially hypoxic areas (Fig. 3b). Research vessels sampled the area from 13 to 20 June 1983. In contrast to 1982, no major phytoplankton blooms were found and only a small area west of Marsh Island was predicted by the CZCS to be potentially hypoxic. The research vessels confirmed the prediction since no hypoxia was found along the entire coast until they neared the satellite predicted area (about 7 days after overpass). It thus appears that the CZCS may be a valuable tool for predicting areas of potential hypoxic bottom water.

In both 1982 and 1983, the research vessels measured hypoxic conditions in the predicted areas up to 10 days after the CZCS overpass. Maximum community plankton respiration rates of up to $8.0 \text{ mg O}_2 \text{ m}^{-3} \text{ h}^{-1}$ have been measured in the bottom waters off the Louisiana coast⁶. At that rate, assuming no mixing, re-aeration or photosynthesis, about 8 days would be required to reduce bottom oxygen concentrations from ambient levels of 4.0 to 2.5 mg l^{-1} . The benthic respiration rate (not measured),

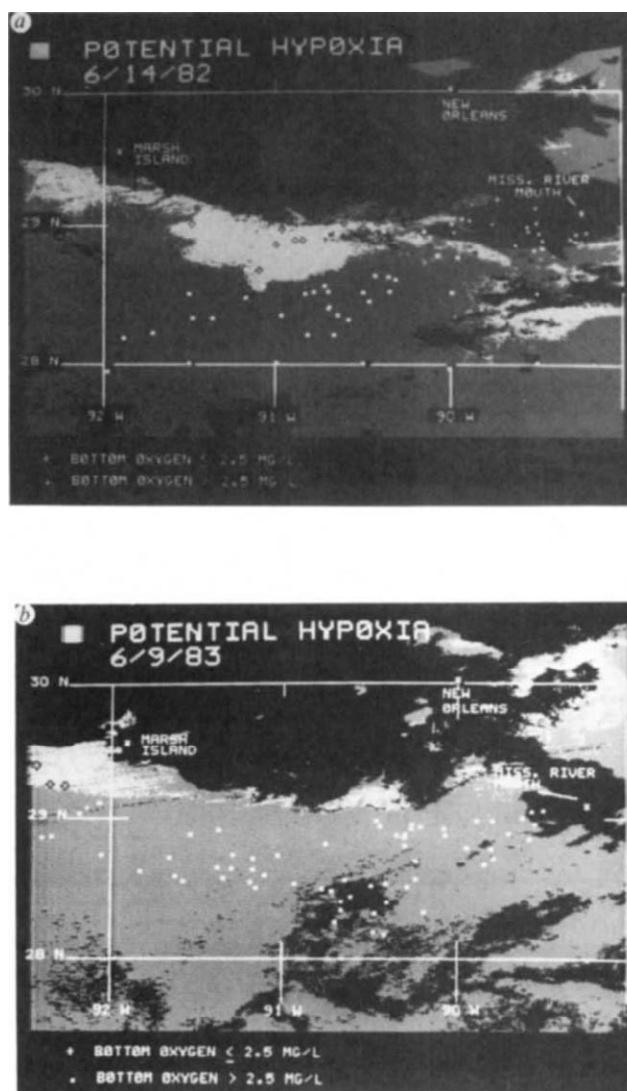


Fig. 3 *a*, Discriminant function classification of 14 June 1982 CZCS image showing area (light grey) of predicted potential hypoxic bottom water. White dots and crosses are vessel locations (15–24 June 1982) with bottom water oxygen concentrations greater and less than or equal to 2.5 mg l^{-1} , respectively. *b*, Discriminant function classification of potential hypoxia on 9 June 1983 using function coefficients from 14 June 1982. Station symbols are the same as *a* (vessels 13–20 June 1983).

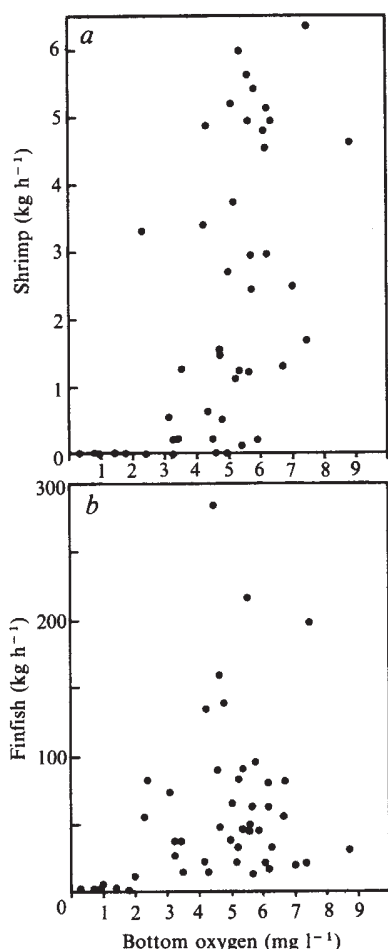


Fig. 4 Shrimp (a) and finfish (b) catch rates plotted against bottom oxygen concentration at research vessel bottom trawl stations for June 1982.

however, may also contribute to increased oxygen depletion rates⁶. The persistence of hypoxic conditions after the decline of surface phytoplankton blooms would depend primarily on vertical and horizontal mixing, advection and photosynthesis in the bottom layer. Thus, if a surface bloom is detected by the CZCS, the potential hypoxia could be predicted to occur up to 10 days later. However, absence of a surface bloom in a particular image does not necessarily mean that hypoxia is absent, due to the unknown time lag associated with persistence. A time series of CZCS images would, however, help to eliminate this uncertainty. Plankton species associated with surface blooms were not measured during the SEAMAP surveys.

The effect of hypoxic bottom waters on living marine resources is not fully understood. Trawl catch rates of shrimp and finfish at the 1982 SEAMAP stations, however, appeared to be greatly influenced by bottom oxygen concentrations (Fig. 4). Above bottom oxygen concentrations of about 2.5 mg l^{-1} , catch rates varied without any apparent dependence on oxygen concentrations. Below this concentration, catch rates were virtually zero. It is not known if the lack of these animals in areas reporting hypoxic conditions was due to direct mortality or to avoidance by the animals. Fish kills in areas of the Gulf of Mexico known to experience hypoxic conditions have, however, been reported^{2,16}.

The ability to detect and monitor hypoxic and potentially hypoxic conditions from space could have significant fisheries implications. If the CZCS were an operational instrument, maps of potentially hypoxic areas could be produced for distribution to the fishermen to be used as strategic fishing guides. Such maps could also be used to plan research programmes on the

effects of hypoxia on the fishing resources. Hypoxia in the northern Gulf normally develops during the summer, the time of peak offshore migration of white shrimp. If further research shows that hypoxia during this stage of the life cycle is a critical factor, the operational mapping of hypoxic areas from space could provide a useful aid in the management of the resource.

Received 16 January; accepted 11 May 1984.

1. Stunz, W. E., Sanders, N., Leming, T. D., Baxter, K. N. & Barazotto, R. M. *Coastal Oceanogr. Climatol. News* 4, 37–38 (1982).
2. Bedinger, C. A. Jr, Childers, R. E., Cooper, J., Kimball, K. T. & Kwok, A. in *Ecological Investigations of Petroleum Production Platforms in the Central Gulf of Mexico*, 1–53 (Rep. BLM Contract AA5551-CT8017 Southwest Research Institute, Texas, 1981).
3. Ragan, J. G., Harris, A. H. & Green, J. H. *Prof. Pap. Ser. Biol.* 3, (Nicholls State Univ., 1978).
4. Turner, R. E. & Allen, R. L. *Contr. mar. Sci.* 25, 161–172 (1982).
5. Boesch, D. F. *Coastal Ocean Pollut. Assess. News* 2, 25–28 (1983).
6. Turner, R. E. & Allen, R. L. *Contr. mar. Sci.* 25, 173–179 (1982).
7. Hovis, W. A. *et al. Science* 210, 60–63 (1980).
8. Gordon, H. R. *et al. Science* 210, 63–66 (1980).
9. Gordon, H. R. & Clark, D. K. *Appl. Opt.* 20, 4175–4180 (1981).
10. Smith, R. C. & Wilson, W. H. in *Oceanography From Space* (ed. Gower, J. F. R.) 281–294 (Plenum, New York, 1981).
11. Clark, D. K. in *Oceanography from Space* (ed. Gower, J. F. R.) 227–237 (Plenum, New York, 1981).
12. Gordon, H. R. *et al. Appl. Opt.* 22, 20–36 (1983).
13. Smith, R. C. & Baker, K. S. *Mar. Biol.* 66, 269–279 (1982).
14. Yentsch, C. S. in *Remote Sensing Applications in Marine Science and Technology* (ed. Cracknell, A. P.) 263–297 (Reidel, Hingham, 1983).
15. Clark, D. K., Baker, E. T. & Strong, A. E. *Boundary-Layer Met.* 18, 287–298 (1980).
16. Harper, D. E., McKinney, L. D., Salzer, R. B. & Case, R. J. *Contr. mar. Sci.* 24, 53–79 (1981).

Greater contribution to blood lead from water than from air

P. C. Elwood, J. E. J. Gallacher, Karin M. Phillips, B. E. Davies* & C. Toothill†

MRC Epidemiology Unit, 4 Richmond Road, CF2 3AS, UK

* Department of Geography, University of Wales, Aberystwyth, Dyfed SY23 3DB, UK

† Supra-regional Assay Service Lead Laboratory, Wharfedale General Hospital, Leeds LS21 2LY, UK

Concern about the levels of lead in blood is widespread. There is uncertainty, however, about the relative importance of the various environmental sources. Lead in petrol is widely assumed to be one of the most important sources and air and dust have been identified as the main routes to man. Water is regarded as an important source in areas with a plumbosolvent water supply, but of little or no importance in other areas. In order to evaluate the contribution to blood lead by various environmental sources, we have conducted surveys of random samples of women in areas of Wales chosen to represent very different levels of exposure to traffic. We report here that lead in air makes a small, but significant, contribution to blood lead but there is no evidence of any contribution from dust. Although in none of the areas were high levels of lead detected in water, water emerges as an important contributor to blood lead.

Our surveys were based on representative samples of women in five areas chosen to represent different exposures to traffic: Cardiff, a major city which has 30,000–40,000 vehicles per day on the major roads; dwellings within 100 m of the elevated motorway in Port Talbot, with around 27,000 vehicles per day on the motorway and around 16,000 on the major roads in the area; dwellings beside major roads in three valleys in South Wales, with around 15,000 vehicles per day; dwellings in culs de sac off these valley roads in which there is no through traffic; and Henllan, a village with about 500 vehicles per day. No area had a plumbosolvent water supply or was near an industrial source of lead or a lead mine. Women were selected because they reflect environmental lead levels more closely than men. Cooperation was high throughout and samples were obtained from around 90% of the women in each area. From the dwellings of a random sub-sample, air samples (4.5 l mm^{-1} for 7 days) were taken at 1 m above floor level in the main living room. The

Table 1 Environmental and venous blood lead levels of representative population samples

	Indoor air lead ($\mu\text{g m}^{-3}$)	Pavement dust lead ($\mu\text{g g}^{-1}$)	Household dust lead ($\mu\text{g g}^{-1}$)	Water lead (mg l^{-1})	Blood lead ($\mu\text{g dl}^{-1}$)
Urban area	0.27 0.06–1.15 (34)*	173 22–1341 (29)*	209 44–998 (51)	0.014 0–0.088 (54)	13.2 $\pm$ 4.76 (56)
Motorway area	0.17 0.04–0.52 (16)*	185 39–879 (18)*	259 62–1075 (29)	0.004 0–0.029 (17)	9.7 $\pm$ 2.69 (31)
Roadside dwelling	0.18 0.05–0.61 (14)*	355 172–731 (42)	200 49–808 (41)	0.007 0–0.041 (39)	8.9 $\pm$ 2.48 (42)
Culs de sac dwellings	0.12 0.03–0.37 (14)*	269 83–870 (28)	177 40–788 (27)	0.004 0–0.056 (25)	8.9 $\pm$ 2.90 (30)
Village	0.04 0.01–0.13 (9)*	—	177 36–851 (31)	0.003 0–0.021 (28)	7.9 $\pm$ 1.86 (33)

Values are geometric mean and total ranges for air lead; means and 95% ranges after log transformation of dust levels, and after cube root transformation of water levels; and mean $\pm$ s.d. for venous blood lead. Number of samples in parentheses.

*Random sub-samples of dwellings.

pavement outside every house ($\sim 2 \text{ m}^2$) was swept and all loose dirt collected, and in every dwelling the contents of the household vacuum cleaner and a 'kettle sample' of water from the cold kitchen tap were taken. Dust samples were dried, sieved (1 mm) and lead extracted by hot nitric acid. Water samples were acidified. Lead estimations were by atomic absorption flame spectrophotometry.

Table 1 shows that in all cases the air lead levels are low but they show marked differences ($P < 0.001$) which are consistent with the traffic flow in the various areas. Although differences in lead levels in pavement dust are statistically significant, they are small relative to the marked differences in traffic. House dust lead levels show very little difference between areas ($P > 0.05$). While these lead levels are lower than those reported by some other workers, methods are not standardized in environmental work and direct comparisons cannot be made. Water lead levels are all very low and only a very few are above levels suggested in the current EC Directive limits (0.05 mg l^{-1})¹. The mean blood lead levels show significant heterogeneity ($P < 0.001$) and increase with increasing traffic levels.

We have examined the data by regression, considering blood lead as dependent on all the environmental sources measured. Clearly results obtained from such an analysis are sensitive to some extent to the transformation used for each variable. We have based decision on transformations on other published studies, and on the fit obtained in our own analyses by different transformations. The first estimate of air lead exposure used for each woman was the mean of all the measurements made in her area of residence. In this model, therefore, it was assumed that differences in air lead measurements made in the various dwellings within a defined area are to a large extent due to sampling and analytical errors, and do not reflect real differences in the exposures of women within that area. Entering the mean air lead levels, together with all the other individual environmental lead levels, generated the following regression:

$$\log \text{ blood Pb} = 1.06 + 0.18 \log \text{ mean air Pb} \\ - 0.02 \log \text{ dust} + 0.62 \text{ water Pb}^{1/3} \quad (1)$$

Regression 'explains' 38% of the variance in blood lead. Pavement dust level does not contribute and the units for the other measurements are as in Table 1. The slope of the blood lead–air lead regression is usually stated for an increase from 1 to $2 \mu\text{g m}^{-3}$ air lead² and entering these levels into equation (1) gives a slope of 1:1.8. Although this represents an extrapolation, it does give a slope which closely agrees with most other authors².

If regression is based on air lead measurements made in individual dwellings, then the following equation is generated:

$$\log \text{ blood Pb} = 0.90 + 0.06 \log \text{ air Pb} \\ + (0.58 \text{ water Pb})^{1/3} \quad (2)$$

House dust no longer contributes and the model 'explains' 23% of the variance in blood lead. Linear versions of these two models 'explain' only marginally less, in each case, of the variance of blood lead levels.

In an attempt to make a realistic evaluation of lead in each source we have entered into the regression equations two levels of lead for a particular source, holding the other sources constant at their overall mean levels. The two levels entered are, in each case, the limits of the 95% range (mean ± 2 s.d.) for that source. Thus, to evaluate water as a source of lead we have entered water lead into the two regressions at 0 and 0.06 mg l^{-1} (representing the 95% range) and held air at $0.16 \mu\text{g m}^{-3}$ and dust at $204 \mu\text{g g}^{-1}$ (representing overall mean levels). In the evaluation of air lead we have used a different pair of levels for each regression: for equation (1) we used two area mean levels, and for equation (2) two individual air lead levels, both pairs representing 95% ranges, the latter derived from the actual range of the untransformed data. On the basis of regression (1), this indicates a range in the contribution from air of up to about $3.1 \mu\text{g dl}^{-1}$, and up to $5.5 \mu\text{g dl}^{-1}$ from water. Using regression (2), the range in contributions from air is up to $1.2 \mu\text{g dl}^{-1}$, and $4.9 \mu\text{g dl}^{-1}$ from water. The contribution of dust is negligible in each regression.

One of the strengths of the data we present is that they are based on representative samples of subjects and their dwellings, within defined areas, and these were chosen to represent 'typical' areas in Wales. Furthermore, the methods for collection of environmental samples, and the analytical methods were identical throughout.

Our data on air lead are not easy to interpret. The use of mean air lead at 'area' level confounds the effects of differences between the areas in social factors, diet and so on, with the estimate of the effect of air. Therefore, equation (1) is certain to give an overestimate and while equation (2) does not eliminate confounding completely, it is probably to be preferred. A further analysis in which an 'area' factor is introduced, as suggested by Azar *et al.*³, gives a very low estimate of the air effect, but the 'area' factor undoubtedly removes some of the air lead effect itself.

The implications of our data on water lead are of great interest. In many studies of this kind, water has been either ignored^{3–5} or dismissed as unimportant⁶. In studies in other areas we have shown that water can be an enormously important source of lead. For example, in an area in North Wales with a plumbosolvent supply the blood lead levels of residents were double⁷. Even in a hard-water area, water can be sufficiently plumbosolvent to raise blood lead levels significantly^{8,9}. The present results indicate that even in areas with quite low water lead levels, water is an important source of lead.

We thank Dr M. Morton of Tenovus, Cardiff for water lead

estimations and the Environmental Health Officers in each area for assistance with air lead measurements.

Received 27 February; accepted 25 May 1984.

1. EC Directive (1977) Council of the European Communities *Off. J. Eur. Commun.* **L105**, 10 (1977).
2. Chamberlain, A. C. *Atmos. Environ.* **17** (4), 677-692 (1983).
3. Azar, A. Snee, R. D. & Habibi, K. *Lead envir. Qual. Saf. Suppl.* **2**, 254-288 (1975).
4. Roels, H. A. *et al. Envir. Res.* **22**, 81-94 (1980).
5. Yankel, A. J., Von Lindern, I. H. & Walter, S. D. *J. Air Pollut. Control Ass.* **27**, 763-767 (1977).
6. Angle, C. R. & McIntire, M. S. *J. Tox. envir. Hlth* **5**, 855-870 (1979).
7. Thomas, H. F., Elwood, P. C., Welsby, E. & St Leger, A. S. *Nature* **282**, 712-713 (1979).
8. Thomas, H. F., Elwood, P. C., Toothill, C. & Morton, M. *Lancet* **i**, 1047-1048 (1981).
9. Elwood, P. C., Phillips, K. M., Lowe, N., Phillips, J. K. & Toothill, C. *Hum. Tox.* **2**, 645-648 (1983).

Increase of corticotropin-releasing factor staining in rat paraventricular nucleus neurones by depletion of hypothalamic adrenaline

Éva Mezey, J. Z. Kiss, L. R. Skirboll, M. Goldstein & J. Axelrod

Laboratory of Cell Biology, Clinical Neuroscience Branch and Laboratory of Clinical Science, National Institute of Mental Health, Bethesda, Maryland 20205, USA

In response to stress, adrenocorticotrophic hormone (ACTH) is released by corticotrophs in the anterior pituitary under the control of several central and peripheral factors^{1,2} including corticotropin-releasing factor (CRF), which was recently isolated from the brain and sequenced³. Immunocytochemical studies have shown that most of the CRF-containing cell bodies that project to the median eminence are present in the hypothalamic paraventricular nucleus (PVN)⁴⁻⁶. A dense PNMT(phenylethanolamine-N-methyltransferase)-containing fibre network was also observed in the same region^{7,8}—PNMT is the final enzyme in the biosynthesis of adrenaline⁹ and has been demonstrated in the brain¹⁰. In the present study we found an association of adrenergic nerve fibres and CRF neurones by immunohistochemistry using antisera to PNMT and CRF. To examine the functional significance of the adrenergic projection to the PVN, we blocked the synthesis of adrenaline using a specific inhibitor of PNMT. The depletion of adrenaline resulted in an increase in CRF immunoreactivity. The present results suggest that, as well as catecholamines which regulate ACTH release at the anterior pituitary level via a β_2 -adrenergic receptor mechanism⁹, central catecholamines (mainly adrenaline) also affect ACTH release through their action on CRF cells. Peripheral catecholamines seem to have a direct stimulatory effect on the pituitary corticotroph cells¹¹, whereas the present findings suggest that central adrenaline-containing neurones have an inhibitory role in the physiological response to stress.

In addition to CRF, it has been suggested that catecholamines acting at central and/or peripheral sites are involved in regulating the release of ACTH^{2,12-15}. In this regard, Seybold *et al.*¹⁶ suggested that catecholamines tonically inhibit vasopressin and neurophysin release from terminals in the external layer of the median eminence; Kiss *et al.* subsequently showed¹⁷ that the origin of these fibres are the CRF-containing cell bodies in the PVN which in response to adrenalectomy also become vasopressin-positive. Recently, we have shown that β_2 -adrenergic receptor agonists have a direct action on ACTH cells in the anterior pituitary¹¹. The present study was based on earlier findings that revealed a dense network of catecholamine-containing fibres in the PVN^{7,8,18} and showed an inverse relationship between adrenaline levels in the brain and corticosterone concentrations in the plasma¹⁴. We wanted to determine whether adrenaline has a direct action on the CRF neurones at a central level to influence ACTH release.

Rats were perfused with 4% paraformaldehyde, then their brains were dissected and processed for indirect fluorescence

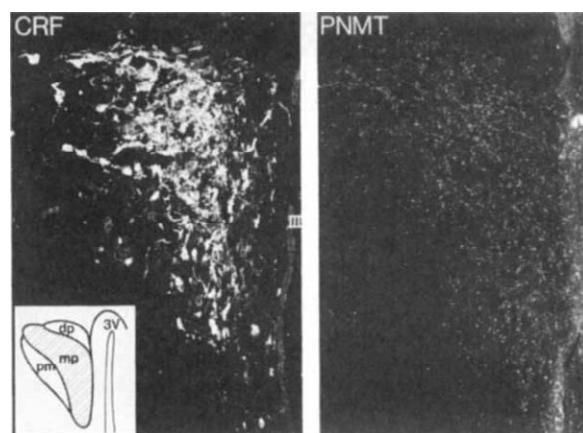


Fig. 1 Immunostaining with anti-CRF and anti-PNMT antisera in the hypothalamic PVN. The insert is a schematic drawing of neurone subpopulations in the PVN at the level presented. The shaded area is the medial parvocellular (mp) region where the CRF-immunopositive cell bodies and the PNMT-positive nerve fibres are located. 3V, Third ventricle; dp, dorsal parvocellular subdivision; pm, posterior magnocellular subdivision. $\times 65$.

Methods: Rats were perfused with 4% paraformaldehyde, then their brains were removed and postfixed in the same fixative for 1 h, after which they were placed in phosphate-buffered saline (PBS) containing 5% sucrose, and washed overnight at 4 °C before processing. Coronal, 4–10 μ m-thick frozen sections were cut in a cryostat, thawed on gelatin-coated slides and incubated with the primary antisera for 24–48 h at 4 °C. The specificities of the anti-CRF and anti-PNMT antisera have been described previously^{7,22}. The antisera were used at a 1:400 dilution in PBS containing 0.6% Triton X-100. The sections were incubated with rhodamine-conjugated secondary antibody at 37 °C for 30 min, washed, mounted with glycerol/PBS and examined by epifluorescence using a Zeiss fluorescent microscope. No staining was observed in the presence of an excess of CRF (10 μ g ml⁻¹ antiserum).

immunocytochemistry (for details see figure legends). In some cases colchicine, an inhibitor of axonal transport, was injected intraventricularly 24 h before perfusion to induce a build-up of CRF peptide in cell bodies. Sections were stained with antisera against PNMT and CRF. On examination of adjacent sections, it was observed that throughout the PVN the distribution of PNMT-immunopositive fibres and CRF-containing neurones overlapped (Fig. 1). Furthermore, double staining of the same sections showed a close juxtaposition of PNMT-containing fibres and CRF-positive cell bodies.

To examine whether adrenaline might affect CRF metabolism, we treated a group of rats with a specific inhibitor of PNMT activity (LY 134046; Lilly Laboratories). As PNMT is the enzyme that converts noradrenaline to adrenaline⁹ and because it has been demonstrated in the rat brain¹⁰, inhibition of this enzyme should reduce the adrenaline content in the brain. It has been reported that animals treated with LY 134046 show a 100% depletion of hypothalamic adrenaline content, but the drug does not affect other brain catecholamines¹⁹. It has also been shown that peripheral catecholamines are not affected by this treatment¹⁹.

Drug-treated and saline-injected (control) animals were perfused and processed for immunostaining using either the indirect fluorescence²⁰ or the avidin-biotin staining procedure^{21,22}. In control animals few, if any, cell bodies were stained with anti-CRF antiserum. This finding is in agreement with the results of other groups who have examined rat brains without injecting the animals with colchicine⁴⁻⁶. On the other hand, numerous CRF-containing cell bodies were visualized in the PVN of animals which had been pretreated with the PNMT inhibitor (Fig. 2a,b). In a separate experiment, rats injected with PNMT inhibitor or saline were also given colchicine 24 h before perfusion. In saline-injected, colchicine-treated animals, the CRF-containing neurones in the medial parvocellular subdivision of the PVN, described by others⁴⁻⁶, were observed. However, rats

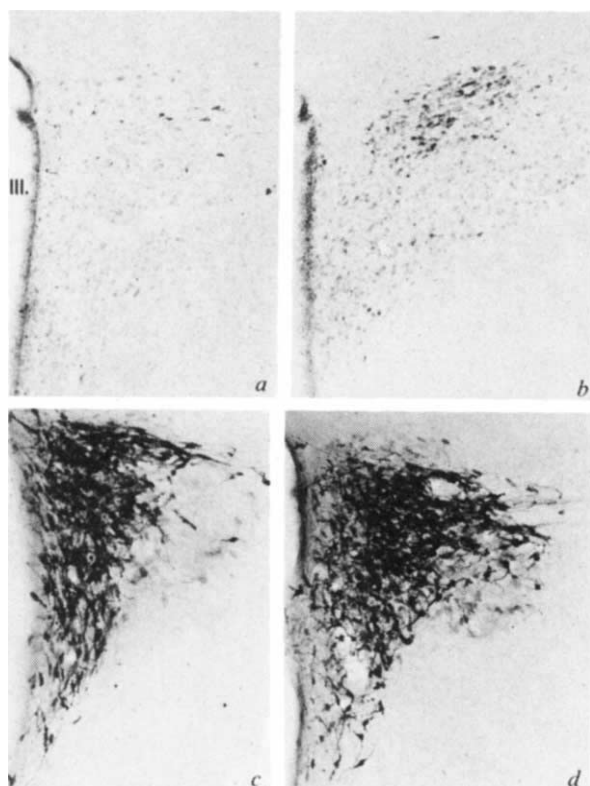


Fig. 2 Immunostaining with anti-CRF antiserum in the PVN. *a*, Saline-injected rat, non-colchicine treatment. Few CRF-positive cell bodies are visible. *b*, Rat injected with PNMT, no colchicine treatment. Note the increase in the number of CRF-positive cells in the medial parvocellular region compared with *a*. *c*, Saline-injected, colchicine-treated rat. The population of CRF cell bodies described elsewhere⁴⁻⁶ can be seen. *d*, Rat injected with PNMT inhibitor, plus colchicine treatment. Note that the combination of PNMT inhibition plus colchicine treatment produces an increase in the number of CRF-positive neurones compared with colchicine alone (*c*). III, Third ventricle. $\times 65$.

Methods: Animals received daily intraperitoneal injections of the PNMT inhibitor (40 mg per kg) for 5 days using physiological saline as the vehicle, or were injected with saline alone (control groups). Colchicine (95 μ g per rat) was injected intraventricularly 24 h before perfusion. Four rats were included in each group. The brains were fixed as described in Fig. 1 legend. Coronal 40 μ m-thick sections were cut in a vibratome and processed according to the ABC (avidin-biotin) method as described elsewhere^{21,22}. The sections were incubated overnight at 4°C in anti-CRF antiserum (1:1,000 in PBS containing 0.6% Triton X-100). After washing the sections in PBS, they were incubated with the biotinylated secondary antibody (1:1,000) at room temperature for 1 h. After rinsing of the sections, they were incubated in the avidin-biotinylated peroxidase complex (1:1,000) at room temperature for 1 h. The peroxidase was then developed using 0.04% 3,3'-diaminobenzidine tetrahydrochloride (Hach) containing 0.025% H₂O. The reaction was stopped after 5 min, then the sections were dehydrated and mounted in Permount.

treated with PNMT inhibitor plus colchicine had almost twice the number of CRF-immunoreactive cells as those injected with colchicine alone (4,440, 4,350 versus 2,490, 2,320 cells).

The increase in CRF staining following PNMT inhibition may be explained in several ways. First, adrenergic fibres may have a tonic inhibitory effect on CRF synthesis and depletion of adrenaline may release this inhibition, causing an increase in CRF production. Such an effect may be due to a direct action of adrenaline on CRF-producing neurones or may be mediated by a group of neurones that also receive adrenergic input and project to the CRF cells in the PVN, thus inducing them to produce more CRF; the first possibility is more likely because adrenaline nerve terminals overlap the CRF neurones. Second,

adrenaline released from terminals in the median eminence may stimulate CRF secretion from nerve endings there and thus PNMT inhibition may cause an accumulation of CRF in the cell bodies due to blockade of release from nerve terminals. This possibility seems unlikely as inhibition of PNMT resulted in enhanced CRF staining even in rats treated with colchicine, which blocks axonal transport and possibly the release of neurotransmitters²³.

We have demonstrated here that CRF-containing cell bodies in the PVN are innervated by adrenergic fibres and that these fibres seem to regulate the synthesis and processing of CRF in the PVN. These observations suggest that in addition to the role of peripheral catecholamines in effecting ACTH release at the pituitary level¹¹, central catecholamines are also involved in regulating ACTH secretion by their action on the hypothalamus. Thus, inhibition of PNMT resulted in a doubling of CRF-positive cells in the PVN of colchicine-treated rats. While the peripheral catecholamines have a stimulatory effect¹¹, our data suggest that centrally acting adrenaline may inhibit ACTH release.

We thank Ray W. Fuller (Lilly Laboratories) for donating the LY 134046, Menek Goldstein (NYU) for the anti-PNMT antiserum and Robert Eskay (NIH) for the anti-CRF antiserum.

Received 23 November 1983; accepted 30 April 1984.

1. Ganong, W. F. *Brain-Endocrine Interaction: Median Eminence: Structure and Function* (eds Knigge, J., Scott, D. S. & Weindl, A.) 254-266 (Karger, Basel, 1972).
2. Jones, M. T. *The Endocrine Hypothalamus* (eds Jeffcoate, S. L. & Hutchinson, J. S. H.) 386-419 (Academic, New York, 1978).
3. Vale, W., Spiess, J., Rivier, C. & Rivier, J. *Science* **213**, 1394-1396 (1981).
4. Bloom, F. E., Battenberg, E. L. F., Rivier, J. & Vale, W. *Regul. Peptides* **4**, 43-48 (1982).
5. Cummings, S., Elde, R., Ellis, J. & Lindall, A. J. *Neuroscience* **3**, 1355-1368 (1983).
6. Swanson, L. W., Sawchenko, P. E., Rivier, J. & Vale, W. *Neuroendocrinology* **36**, 165-186 (1983).
7. Hokfelt, T., Fuxe, K., Goldstein, M. & Johansson, O. *Brain Res.* **66**, 235-251 (1974).
8. Swanson, L. et al. *J. comp. Neurol.* **196**, 271-285 (1981).
9. Axelrod, J. *J. biol. Chem.* **237**, 1657-1660 (1962).
10. Saavedra, J., Palkovits, M., Brownstein, M. J. & Axelrod, J. *Nature* **248**, 695-696 (1974).
11. Mezey, E., Reisine, T., Palkovits, M., Brownstein, M. J. & Axelrod, J. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6728-6731 (1983).
12. Ganong, W. F. *Fedn Proc.* **39**, 2923-2930 (1980).
13. Weiner, R. & Ganong, W. F. *Phys. Rev.* **58**, 905-976 (1978).
14. Roth, K. A. et al. *Life Sci.* **28**, 2389-2394 (1981).
15. Turner, B. B., Katz, R. J., Roth, K. A. & Carroll, B. J. *Brain Res.* **153**, 419-422 (1978).
16. Seybold, V., Elder, R. & Hokfelt, T. *Endocrinology* **108**, 1803-1809 (1983).
17. Kiss, J. Z., Mezey, E. & Skirboll, L. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
18. Swanson, L. W. & Hartman, B. K. *J. comp. Neurol.* **163**, 467-506 (1975).
19. Fuller, R. W. et al. *Biochem. Pharmacol.* **30**, 1345-1352 (1981).
20. Sternberger, L. *Immunocytochemistry* (Wiley, New York, 1979).
21. Hsu Su-Ming & Ree, H. J. *Am J. clin. Path.* **74**, 32-40 (1980).
22. Hsu Su-Ming, Raine, L. & Fanger, H. J. *Histochem. Cytochem.* **29**, 577-580 (1981).
23. Inoue, S. & Sato, H. *J. gen. Physiol.* **50**, 259-288 (1967).

Topological distribution of different forms of neural cell adhesion molecule in the developing chick visual system

Burkhard Schlosshauer, Uli Schwarz & Urs Rutishauser*

Max-Planck-Institut für Virusforschung, 7400 Tübingen, FRG

* Department of Developmental Genetics and Anatomy, School of Medicine, Case Western Reserve University, Cleveland, Ohio 44106, USA

The cell adhesion molecule isolated from neural tissue (N-CAM) is a membrane glycoprotein which is directly involved in calcium-independent adhesion among nerve cells and their processes (for review see refs 1, 2). N-CAM has an unusual carbohydrate moiety containing a large and variable amount of sialic acid, the variation reflecting both the type of tissue and its developmental age^{3,4}. N-CAM is believed to be a ligand in the formation of cell-cell bonds⁵ and a decrease in sialic acid content from 30% to 10% is associated with a marked enhancement of the molecule's binding activity⁶⁻⁸. Antibodies to N-CAM block its function and inhibit or alter bundling of nerve fibres⁹, retinal cell development¹⁰⁻¹²

and nerve-muscle interaction^{13,14}. Here we use micro-gel electrophoresis¹⁵ to compare N-CAM from several parts of the developing chick visual system. The results indicate that N-CAM from the retina of 5–10-day-old embryos already exists in a relatively sialic acid-poor form, whereas the tectum and optic nerve beyond the eye contain sialic acid-rich N-CAM until much later in development. These studies suggest that the perikaryon and proximal axon shaft of retinoganglion cells have N-CAM with a lower sialic acid content than the distal portion of the axons, and that resulting differences in neurite adhesivity may be an important factor in the formation of the optic system.

In the chick, retinoganglion axons first leave the eye at about embryonic day 3, gradually building the optic nerve. The first axons reach the tectum at day 6 and cover the entire tectal surface by day 12 (ref. 16). The N-CAM of the optic system on embryonic day 10 was analysed by electrophoresis of SDS-solubilized tissues in micro-polyacrylamide gels (SDS-PAGE)¹⁵, transfer of the proteins to nitrocellulose sheets¹⁷ and staining with affinity-purified rabbit anti-chicken N-CAM (Fig. 1). This experiment revealed that the N-CAM from tissue slices representing several consecutive layers of the retina, including nuclear and fibre layers, has a high electrophoretic mobility characteristic³ of the sialic acid-poor molecule. In contrast, dissected optic nerve and chiasm, optic tract and tectum yielded N-CAM with the lower mobilities associated with the sialic acid-rich form of the molecule. In agreement with previous studies³, removal of sialic acid from tectal N-CAM by boiling yielded material with an SDS-PAGE band pattern similar to that of retinal N-CAM. Addition of protease or neuraminidase inhibitors (Trasylol, trypsin inhibitor, phenylmethylsulphonyl fluoride and 2,3-dehydro-2-deoxy-*N*-acetylneuraminic acid) did not change the banding pattern of the samples, suggesting that the observed differences were not caused by enzymatic degradation during tissue solubilization.

N-CAM from retina, optic nerve and tectum was also analysed for an age-dependent decrease in sialic acid content (Fig. 2). Optic nerve and tectum gave results similar to those previously obtained for whole brain and cerebellum^{4,18}, that is, a progressive developmental increase (in chick, from embryonic day 10 to postnatal day 3) in the amount of faster-migrating components on SDS-PAGE. In contrast, retinal N-CAM not only had a high mobility on SDS-PAGE at embryonic days 10–17, but already exhibited this sialic acid-poor form as early as day 4. Therefore, the developmental regulation of the sialic acid content of retinal N-CAM seems to be distinct from most other central nervous tissues.

Retinoganglion cells contribute to the N-CAM detected in retinal, optic nerve and tectal tissues. For the optic nerve, this contribution is substantial, being diluted out only by non-neuronal cells. In the retina, other neuronal cell types are also included, but the similarity of the N-CAM obtained from the different layers of this tissue (Fig. 1) suggests that the sialic acid-poor form is expressed throughout. Therefore, during embryonic days 5–10, two retinal neurite populations can be distinguished: (1) the plexiform layers and the optic fibre layer within the retina which express sialic acid-poor N-CAM and (2) the axons projecting out of the retina which express sialic acid-rich N-CAM. This conclusion implies that different forms of N-CAM exist on different regions of the same retinoganglion cell. That is, the perikaryon and intra-retinal axon have N-CAM with relatively less sialic acid than the part of the axon lying beyond the optic stalk.

The possibility of N-CAM heterogeneity on the same cell has been explored further by analysis of the neurites which emerge from retinal tissue slices in culture¹⁹. These processes, which have been identified as ganglion cell axons²⁰, can be dissected away from the explant and are not contaminated with myelin or glial cells. Figure 3 shows two-dimensional gels of N-CAM from the axons and from explants containing perikarya. In accord with our studies on optic nerve, the axons had N-CAM with the lower mobility and more acidic isoelectric point characteristic³ of the sialic acid-rich form, while the explant yielded

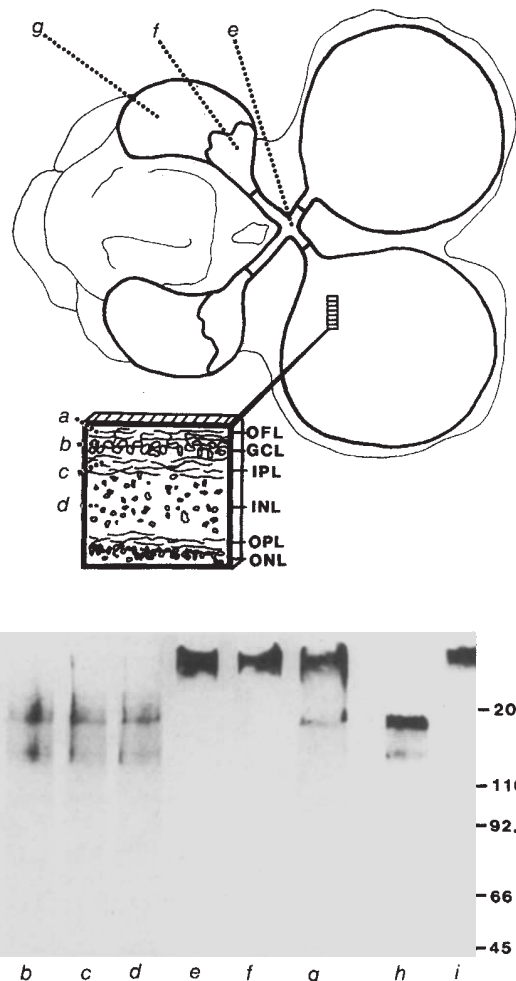


Fig. 1 Electrophoretic comparison of N-CAM from different parts of the chick visual system at embryonic day 10. Retinal layers (insert: *a–d*) were obtained by sectioning of frozen, flat-mounted 10-mm² pieces of neural retina from the indicated region about midway between the centre and margin of the retina. In this region, the optic fibre layer (OFL), ganglion cell layer (GCL), inner plexiform layer (IPL) and inner nuclear layer (INL) were approximately 30, 60, 30 and 120 μ m thick, respectively. Successive 15- μ m sections were cut and examined for uniformity of thickness to insure against buckling of the tissue. The fractions *a–d* represented sections 1–2, 3–6, 7–8 and 9–14. Contamination of one layer by its neighbour would be less than 30% and did not extend beyond adjacent layers. Thus, fraction *a* is derived from ganglion cell axons (and possibly some ganglion cell soma) from within the eyecup and is not significantly contaminated either by other neurones of the retina or by regions of ganglion cell axons lying beyond the optic fissure. The chiasma (*e* in the schematic drawing of the visual system) was dissected with stumps of the optic nerve lying just beyond the optic stalk, and with the proximal tract; the distal optic tract (*f*) was peeled off from the tectal surface (*g*). These samples were solubilized in SDS sample buffer without boiling and ~500 ng total protein per lane was fractionated by SDS-PAGE on microgels¹⁵. The fractionated proteins were transferred to nitrocellulose sheets¹⁷, reacted with affinity-purified rabbit anti-(N-CAM) followed by horseradish peroxidase-conjugated goat anti-rabbit IgG, and finally stained by treatment with 4-chloro-1-naphthol²¹. All retinal layers including the fibre layers (*a* and *c*) contained N-CAM with the high mobility characteristic of the sialic acid-poor molecule. In contrast, at this developmental stage, all other parts of the visual system (*e–g*) primarily contained N-CAM with a high content of sialic acid. The presence of sialic acid-rich N-CAM in *a–d* would have been easily detected at even 1/20th of levels shown in *e–g*. Boiling of the tectum sample for 30 min to release sialic acid (*h*) yielded the higher mobility form of the molecule. Forebrain N-CAM (*i*) was found to be similar to optic nerve and tectal N-CAM. Other abbreviations: OPL, outer plexiform layer; ONL, outer nuclear layer. Molecular weight standards are given on the right of the gels ($\times 10^{-3}$).

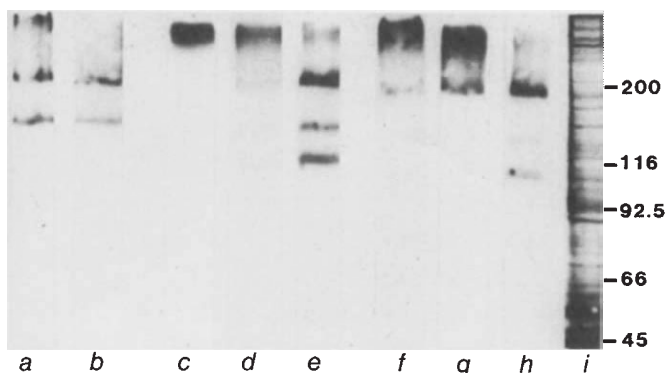
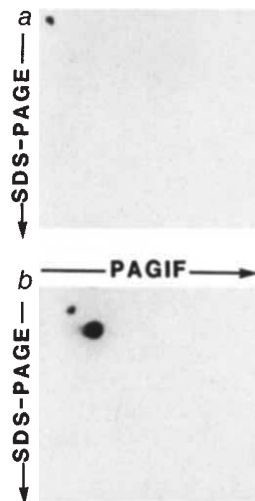


Fig. 2 Modification of N-CAM form during development of the visual system. SDS-PAGE of N-CAM from retina on embryonic day 5 (a) and day 10 (b); from optic chiasma on embryonic day 10 (c), day 18 (d), and on day 3 after hatching (e); and from tectum on embryonic day 10 (f), day 18 (g) and day 3 after hatching (h). Lane i is a silver-stained gel of total tectal protein. Note that in each tissue there is a gradual change from slow- to fast-migrating forms of N-CAM during development, but in retina the sialic acid-poor forms appear much earlier and are already present at day 5. Molecular weight standards are given on the right ($\times 10^{-3}$).

Fig. 3 Two-dimensional separation of N-CAM from cultured explant axons derived from retinoganglion cells²⁰ (a) and from explant tissue containing perikarya (b). Two-dimensional separation of samples on polyacrylamide microgels by isoelectric focusing (PAGIF) and electrophoresis in SDS (SDS-PAGE) was performed as described previously¹⁵. The acidic end of the pH gradient on the gels is on the left, high-molecular weight proteins are at the top of each gel. The pure fraction of cultured retinoganglion cell axons obtained by dissection contains only acidic N-CAM with a low mobility, while the perikarya-rich explant contains primarily faster-migrating forms of N-CAM with a more neutral isoelectric point. Cross-contamination of the two fractions would have been detected at 5% of the amounts shown here.



molecules with the higher mobility and more neutral isoelectric point of N-CAM isolated from retinal tissue.

The primary mechanisms by which N-CAM might appear in different forms on distinct regions of the same cell are the synthesis of different molecules followed by their insertion onto the appropriate membrane site, and modification of N-CAM after its integration into the membrane. At this point it is impossible to choose between these two possibilities. The latter might seem more attractive in that in the embryo the 'jump' in sialic acid content occurs at or just beyond a natural anatomical boundary composed of the eyecup and optic stalk. Therefore, the presence of an extracellular neuraminidase-like activity within the eye could provide a simple explanation for our *in vivo* results. However, the different forms of N-CAM were also observed with cultured retinal explants, and this result would require that the putative sialidase be produced by the explant and act only locally.

The present studies thus indicate that all parts of the visual system express the same N-CAM polypeptide but that inside the retina the N-CAM contains a carbohydrate moiety with considerably less sialic acid than that of the N-CAM outside the retina. This change does not seem to be graded along the

visual pathway, but instead occurs in a single step at or near the optic stalk. Assuming that these differences in sialic acid content are associated with a change in N-CAM binding affinity⁶⁻⁸, it is interesting to speculate that the stronger binding of retinal N-CAM reflects the need for stable fasciculation during formation of the optic nerve. This hypothesis is supported by the recent demonstration that a reduction in adhesion by injection of anti-(N-CAM) Fab into the eyecup of 4-day-old chick embryos disrupts normal fibre order at the optic fissure¹². Once the optic nerve leaves the eye, it interdigitates with the contralateral nerve and spreads out over the tectal surface. These events involve some remodelling of fascicles and possibly axon-tectal contacts, which might require the expression of an N-CAM with a lower binding affinity. At later stages of development, when most retinal fibres have reached their target tissue, the N-CAM in optic nerve and tectum also loses part of its sialic acid. This shift might indicate that once the visual projection is established, plasticity is no longer desirable and therefore the higher-affinity N-CAM is expressed.

This work was supported in part by USPHS grant HD-18369. We thank S. Deiss for technical assistance and I. Baxivanelis and M. E. Bleakley for typing the manuscript.

Received 20 December 1983; accepted 8 May 1984.

1. Rutishauser, U. *Cold Spring Harb. Symp. quant. Biol.* **48**, 501-514 (1983).
2. Edelman, G. M., Hoffman, S. & Cunningham, B. A. *Cold Spring Harb. Symp. quant. Biol.* **48**, 515-526 (1983).
3. Hoffman, S. *et al. J. biol. Chem.* **257**, 7720-7729 (1982).
4. Rothbard, J. B., Brackenbury, R., Cunningham, B. A. & Edelman, G. M. *J. biol. Chem.* **257**, 11064-11069 (1982).
5. Rutishauser, U., Hoffman, S. & Edelman, G. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 685-689 (1982).
6. Cunningham, B. A., Hoffman, S., Rutishauser, U., Hemperly, J. J. & Edelman, G. M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3116-3120 (1983).
7. Hoffman, S. & Edelman, G. M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5762-5766 (1983).
8. Sadoul, R., Hirn, M., Deagostini-Bazin, H., Rougon, G. & Goridis, C. *Nature* **304**, 347-349 (1983).
9. Rutishauser, U., Gall, W. E. & Edelman, G. M. *J. Cell Biol.* **79**, 382-393 (1978).
10. Rutishauser, U., Thiery, J.-P., Brackenbury, R. & Edelman, G. M. *J. Cell Biol.* **79**, 371-381 (1978).
11. Buskirk, D. R., Thiery, J.-P., Rutishauser, U. & Edelman, G. M. *Nature* **285**, 488-489 (1980).
12. Thanos, S., Bonhoeffer, F. & Rutishauser, U. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1906-1910 (1984).
13. Grumet, M., Rutishauser, U. & Edelman, G. M. *Nature* **295**, 693-695 (1982).
14. Rutishauser, U., Grumet, M. & Edelman, G. M. *J. Cell Biol.* **97**, 145-152 (1983).
15. Neukirchen, R. O., Schlosshauer, B., Baars, S., Jackle, H. & Schwarz, U. *J. biol. Chem.* **257**, 15229-15234 (1982).
16. Rager, G. H. *Development of Retinotectal Projection in the Chicken* (Springer, Berlin, 1980).
17. Towbin, H., Staehlin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350-4354 (1979).
18. Edelman, G. M. & Choung, C.-M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7036-7040 (1982).
19. Halfter, W., Claviez, M. & Schwarz, U. *Nature* **292**, 67-70 (1981).
20. Halfter, W., Newgreen, D. F., Sauter, J. & Schwarz, U. *Dev. Biol.* **95**, 56-64 (1983).
21. Hawkes, R., Niday, E. & Gordon, J. *Analyt. Biochem.* **119**, 142-147 (1982).

Different forms of p53 detected by monoclonal antibodies in non-dividing and dividing lymphocytes

Jo Milner

Division of Virology, Department of Pathology, University of Cambridge, Cambridge CB2 1QP, UK

The commitment of non-dividing cells (in G_0) to enter division can be studied using primary cultures of lymphocytes. The cells are stable in G_0 unless stimulated by mitogen such as concanavalin A (Con A). Commitment to enter the division cycle depends on the expression of gene(s) induced by Con A^{1,2} and the synthesis of p53 protein correlates with this commitment step³. I show here that in unstimulated cells, a second form of p53 is synthesized and is restricted to G_0 .

The possibility that p53 may function in normal cell division is indicated by the fact that its synthesis in lymphocytes can be detected after, but not before, mitogenic stimulation^{3,4}. In these earlier experiments^{3,5} ³⁵S-labelled p53 was precipitated using anti-p53 antiserum (from mice bearing simian virus 40 (SV40)-induced tumours) or an anti-p53 monoclonal antibody

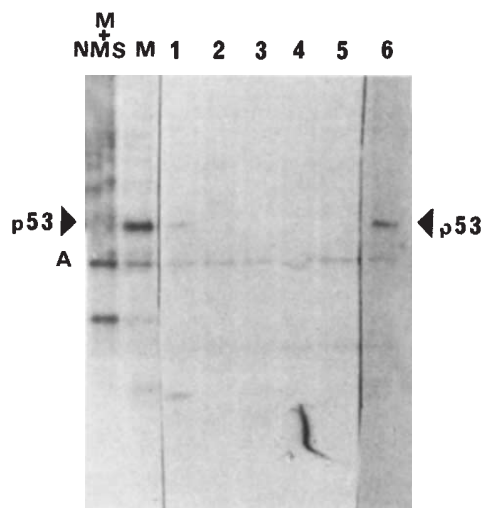


Fig. 1 p53 synthesis in unstimulated lymphocytes (in G_0). Splenic lymphocytes from BALB/c mice were maintained in culture (5×10^6 cells per ml RPMI medium) without serum for 20 h. The medium was then replaced with methionine-free RPMI plus ^{35}S -methionine ($1,000 \text{ Ci mmol}^{-1}$; 0.25 mCi ml^{-1}) and the cells were labelled for 10 min (no further label was incorporated into immunoprecipitable p53 with longer labelling, consistent with a short half life for p53). Immediately after labelling the cells were lysed⁴ and equal aliquots of lysate (containing equal trichloroacetic acid-precipitable ^{35}S counts) were immunoprecipitated with different monoclonal antibodies to p53 as follows: track 1, PAb248; 2, PAb242; 3, PAb421; 4, PAb122; 5, 200.47; 6, RA3.2C2. All the monoclonal antibodies precipitated p53 from SVA31E7 cells. p53 from spontaneously transformed 3T3 cells (kindly donated by David Lane) was used as marker (track M); the cells were labelled with ^{35}S -methionine and lysate immunoprecipitated with RA3.2C2 (track M; the same result was observed with PAb421) or with normal mouse serum (track M+NMS). p53 was visualized by autoradiography (fluorography) after electrophoresis on a 15% polyacrylamide gel and the gel was exposed to X-ray film for 28 days at -70°C .

(PAb421). Unstimulated and stimulated cells were compared on the basis of equivalent cell numbers and equivalent ^{35}S -labelled protein content: no p53 was detected in the unstimulated cells. The first detectable synthesis of p53 was found to coincide with mitogenic commitment and its induction to be, directly or indirectly, under transcriptional control³.

In the present study p53 synthesis was examined using the following panel of monoclonal antibodies to p53: PAb421 (ref. 5), PAb122 (ref. 6), PAb242 (ref. 7), RA3.2C2 (ref. 8), PAb248 (ref. 7) and 200.47 (ref. 9). As expected from earlier results, PAb421, PAb122, PAb242 and 200.47 failed to detect p53 in lysate of unstimulated cells. However, both PAb248 and RA3.2C2 precipitated p53 from aliquots of the same lysate of unstimulated cells (Fig. 1). This result was unexpected and led to a detailed re-examination of p53 synthesis in stimulated cells. PAb421 precipitated p53 from the stimulated cells, confirming previous observations³; of the other monoclonals only PAb122 detected p53 from the stimulated cells (Fig. 2). Each monoclonal used is known to precipitate p53 from SV40-transformed SVA31E7 cells (results not shown), so failure to bind p53 is not due to loss of antibody-binding capacity. Thus PAb248 and RA3.2C2 specifically detect p53 in unstimulated, but not stimulated, lymphocytes; conversely, PAb421 and PAb122 (and mouse anti-p53 antiserum⁴) detect p53 only after mitogenic stimulation. This discrimination appears to be real, rather than an artefact due to cross-reactivity, because each anti-p53 monoclonal binds a 53,000-molecular weight protein from either unstimulated or stimulated cells, but not from both, as would be expected for a cross-reacting protein.

Having found two immunologically distinct forms of p53, hereafter termed p53₂₄₈, recognized by PAb248 and RA3.2C2,

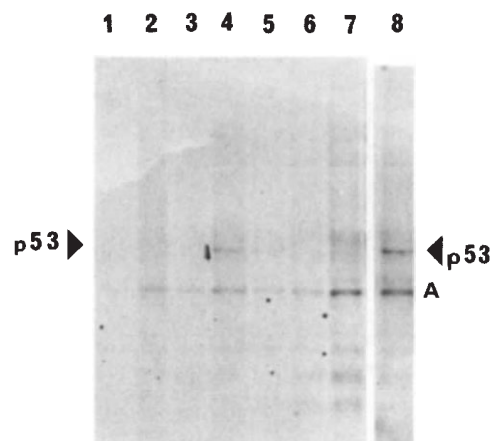


Fig. 2 p53 synthesis in lymphocytes after mitogenic stimulation. Cells were cultured in the presence of Con A ($1 \mu\text{g ml}^{-1}$) for 24 h, otherwise as described in Fig. 1 legend. After 10 min labelling with ^{35}S -methionine the cells were lysed and equal aliquots immunoprecipitated as follows: track 1, with normal mouse serum; 2, RA3.2C2; 3, PAb248; 4, PAb421; 5, PAb122; 6, PAb242; 7, 200.47. p53 from spontaneously transformed 3T3 cells was used as marker (track not shown). Track 8 is from a mixture of p53₂₄₈ and p53₄₂₁ (from different gels). Tracks 1-7 were exposed to X-ray film for 28 days; track 8 for 20 days.

and p53₄₂₁, recognized by PAb421 and PAb122, the synthesis of p53₂₄₈ and of p53₄₂₁ was monitored following mitogenic stimulation. After 2 h stimulation p53₂₄₈ was scarcely detectable and from 4 h to 48 h p53₄₂₁ was the only form detectable (Table 1). No overlap in the synthesis of the two forms of p53 was observed. The timing of changeover from p53₂₄₈ to p53₄₂₁ shows a striking correlation with the time required for lymphocytes to generate intracellular signal(s) necessary for mitogenic commitment^{1,2}.

There are three possible explanations for these results. First, the change in form of p53 may represent a simple conformational change in response to the complex ionic and metabolic events known to occur following mitogenic stimulation of lymphocytes (see, for example, ref. 10). Second, the two forms of p53 may result from post-translational modification(s), with or without coincident conformational effect. Both these possibilities imply the existence of a single p53 protein, the form of which depends on the dividing state of the cell. A third possibility is that p53₂₄₈ and p53₄₂₁ represent two distinct proteins inseparable by polyacrylamide electrophoresis (see Fig. 2, track 8); this possibility could not be further investigated by partial peptide mapping because of the very low amounts of p53 obtained in this system. Whichever explanation proves correct, the important points of the present observations are: (1) two immunologically distinct forms of p53 exist in lymphocytes; (2) the synthesis of each form is restricted to non-dividing cells (for p53₂₄₈) or to dividing cells (for p53₄₂₁); and (3) p53₂₄₈ is replaced by p53₄₂₁ as cells become committed to enter division.

Our findings that the epitopes recognized by RA3.2C2 and PAb421 are not expressed simultaneously on p53 in lymphocytes explains our inability to detect lymphocyte p53 in a non-competitive immunoassay using RA3.2C2 and PAb421 (collaborative unpublished studies with Lionel Crawford and with David Lane). As this same assay detects p53 from SV40-transformed SVA31E7 cells¹¹, the implication is that p53 in SV40-transformed cells may differ from lymphocyte p53 in having both epitopes expressed simultaneously. Alternatively, the SV40 large-T antigen, which binds to and stabilizes p53 (refs 12, 13), may bind both forms of p53 and this would account for the development of antibodies against each form when SV40-transformed cells are used to immunize mice⁵⁻⁷. Syngeneic immunization⁴⁻⁶ appears to elicit antibodies against p53₄₂₁ (that is, PAb421, PAb122 and anti-p53 antiserum) but not against p53₂₄₈,

Table 1 Synthesis of p53₂₄₈ and p53₄₂₁ in lymphocytes, before and at different times after stimulation with Con A

Time after Con A stimulation (h)	p53 synthesis detected using		DNA synthesis
	PAb248	PAb421	
0	+++	0	2,040
2	+	0	1,960
4	0	+	1,840
6	0	++	1,780
8	0	++	1,890
20	0	+++	20,400
42	0	+	85,500

Cells were labelled for 10 min, as described in Fig. 1 legend, and equal aliquots of ³⁵S-methionine-labelled cell lysates were immunoprecipitated separately with the monoclonal antibodies PAb248 and PAb421. ³⁵S-methionine-labelled p53 was visualized by electrophoresis and fluorography (see Fig. 1 legend) and the relative intensities of the p53 bands were assessed by scanning the autoradiographs on a Beckman DU-8 spectrophotometer, using actin as an internal reference band in each track. As the overall background level varied at different times after stimulation, the intensity of p53 is expressed relative to that of actin, on the basis of peak heights, as follows: 0, p53 undetectable; +, 0.5; ++, 0.5–1.0; +++, 1.2–3.4. DNA synthesis is c.p.m. of ³H-thymidine (5 Ci mmol⁻¹; 20 µCi ml⁻¹) incorporated into DNA² per 5 × 10⁵ cells, x, n = 5.

which is detected by PAb248 and RA3.2C2, products of allogeneic⁷ and xenogeneic⁸ immunizations respectively—this raises the possibility that p53₄₂₁ is the autoantigenic form of p53. Cross-blocking studies have shown that PAb248 and RA3.2C2 completely block each other's binding to p53 from SVA31E7 cells, as do PAb421 and PAb122, whereas there is no blocking between PAb248 and RA3.2C2 with PAb421 and PAb122 (refs 7, 11). Thus, the two pairs of monoclonals, PAb248 plus RA3.2C2, and PAb421 plus PAb122, fall into the same pairing on the basis of (1) binding to p53 from unstimulated or stimulated lymphocytes, (2) their initial derivation via syngeneic or non-syngeneic immunization, and (3) their blocking characteristics on p53 from SVA31E7 cells. The lack of binding of 200.47 and PAb242 to lymphocyte p53 is puzzling. Monoclonal 200.47 was developed after immunizing a (C57BL/6 × BALB/c) F₁ mouse with BALB/c methylcholanthrene-induced sarcoma (CMS4) cells⁹ and PAb242 was developed following syngeneic immunization of a BALB/c mouse with SV40-transformed cells⁷: if p53₂₄₈ and p53₄₂₁ form a complex in transformed cells it is possible that 200.47 and PAb242 only bind the complexed form, p53₂₄₈₊₄₂₁. Certainly the epitopes for 200.47 and PAb242 map independently and separately from PAb248 plus RA3.2C2 and PAb421 plus PAb122 (refs 7, 11).

The detection of two immunologically distinct forms of p53 raises the possibility that p53₂₄₈ and p53₄₂₁ represent different functional forms of the protein. The switch in synthesis from p53₂₄₈ to p53₄₂₁ at the time of mitogenic commitment suggests a tight link with the control of division in normal cells.

This work was supported by a CRC project grant. I thank Ed Harlow and Lionel Crawford for kindly donating samples of PAb421, 200.47, RA3.2C2 and PAb122; David Lane for PAb248 and PAb242; and Su Metcalfe for constructive criticism.

Received 22 November 1983; accepted 24 May 1984.

- Milner, S. M. *Nature* **268**, 441–442 (1977).
- Milner, J. *Nature* **275**, 660–661 (1978).
- Milner, J. & Milner, S. *Virology* **112**, 785–788 (1981).
- Milner, J. & McCormick, F. *Cell Biol. Int. Rep.* **4**, 663–667 (1980).
- Harlow, E., Crawford, L. V., Pim, D. C. & Williamson, N. M. *Virology* **39**, 681–689 (1981).
- Gurney, E. G., Harrison, R. O. & Fenno, J. *J. Virol.* **34**, 752–753 (1980).
- Yewdell, J., Gannon, J. & Lane, D. P. *EMBO J.* (submitted).
- Coffman, R. L. & Weissman, I. L. *J. exp. Med.* **153**, 269–279 (1981).
- Dippold, W. G., Jay, G., DeLeo, A. B., Khoury, G. & Old, L. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1695–1699 (1981).
- Ling, N. R. & Kay, J. E. *Lymphocyte Stimulation* (North-Holland, Amsterdam, 1975).
- Benchimol, S., Pim, D. & Crawford, L. V. *EMBO J.* **1**, 1055–1062 (1982).
- McCormick, F. & Harlow, E. *J. Virol.* **34**, 213–224 (1980).
- Oren, M., Maltzman, W. & Levine, A. J. *Molec. cell. Biol.* **2**, 101–110 (1981).

Immortalized B lymphocytes produce B-cell growth factor

John Gordon, Steven C. Ley, Mark D. Melamed, Leonard S. English* & Nevin C. Hughes-Jones

Mechanisms in Tumour Immunity, MRC Centre, Cambridge CB2 2QH, UK

Department of Immunology, Institute of Animal Physiology, Babraham, Cambridge CB2 4AT, UK

* Microbiology Department, School of Medicine, East Carolina University, Greenville, North Carolina 27834, USA

The activation, clonal expansion and terminal differentiation of small resting B lymphocytes primed by an antigen (or antibodies to its receptors) appear to follow an orderly developmental sequence triggered at each stage by distinct soluble cytokines, primarily produced by T lymphocytes^{1,2}. For man, the only known B-cell mitogen independent of accessory cells for its action is the Epstein-Barr virus (EBV). Lymphocytes transformed by EBV are released from the usual constraints on B-cell growth, proliferating continuously in the absence of any exogenous cytokine³. The resultant cell lines are of special interest as they possess certain features compatible with a preneoplastic state of Burkitt's lymphoma, one of two human cancers with which the virus is intimately associated³. We report here that following EBV-transformation, B lymphoblasts release a soluble factor which mimics the B-cell stimulatory product(s) of mitogen-conditioned T lymphocytes. Furthermore, the virally-transformed cells utilize this activity to sustain their own growth. The ectopic production of an otherwise normal growth factor may represent a critical event in the malignant evolution of human lymphomas harbouring the EBV genome.

B-cell stimulatory activity was assayed by the ability of conditioned media to synergize with anti-immunoglobulin (Ig) to induce DNA synthesis in tonsillar B lymphocytes. Initial experiments (not detailed) using medium conditioned by EBV-transformed lymphoblastoid cells in their log phase of growth (LCL-CM) gave variable low or negative responses using a F(ab')₂ rabbit preparation with broad immunoglobulin reactivity (IgG heavy and light chains; Miles) as the synergistic antibody. As immunoglobulins released into culture by the lymphoblastoid

Table 1 B-cell stimulating activity in medium conditioned by B-lymphoblastoid cells

Conditioned medium	Isotype specificity	c.p.m. ³ H-thymidine incorporated into B cells*			
		High density			Low density control
		Control	anti-μ†	anti-γ	
Control	—	220	4,801	2,013	1,494
T cell	—	354	10,432	6,044	5,323
LCL _{TL}	IgM	362	670	4,760	4,846
LCL _{D6}	IgG	249	8,745	252	5,724

Cells were cultured at 10⁵ per well in sterile flat-bottom (growth area 0.32 cm² per well) microtitre plates of ELISA grade (Dynatech) in a final volume of 150 µl for 72 h then labelled with 0.5 µCi ³H-thymidine for the final 16 h. Standard deviations of quadruplicate determination were routinely <10%. B cells were isolated from tonsils by negative selection²⁰. E-rosette positive cells were removed on Ficoll gradients. The non-rosetting population was routinely >97% surface Ig⁺ by immunofluorescence and contained <1% E-rosette positive cells. The high density population represents B cells banding on a discontinuous Percoll (Pharmacia) gradient at layers >60%. Low density B cells were collected from layers <50% Percoll. Microwells were coated with 1 µg IgG from rabbit antisera specific for human μ or γ chains (Dako) by overnight incubation at 37 °C in 100 µl sodium carbonate buffer (pH 9.6) and washed free of unbound protein before use (P. Áman, J. G. and G. Klein, manuscript in preparation). T-cell supernatants were collected from 24 h cultures of tonsillar E-rosette positive cells previously pulsed for 48 h with 1% pokeweed mitogen (Gibco). All conditioned media were added to 33% at initiation of culture. The establishment, cloning and maintenance of EBV-transformed lymphoblastoid cell lines (LCL) have been described elsewhere²⁰. Conditioned medium was collected from cells which had been taken in their log phase of growth and reseeded at a concentration of 3 × 10⁵ per ml for 18 h in fresh medium (RPMI 1640 supplemented with 10% fetal calf serum, 200 mM L-glutamine, 50 µg ml⁻¹ penicillin/streptomycin and 5 × 10⁻⁵ M 2-mercaptoethanol). All conditioned media were passed through a 0.45-µm filter and stored at 4 °C before use on the same day.

Table 2 B-cell stimulating activity from different lymphoid cell lines

B-cell preparation	c.p.m. ^3H -thymidine incorporation in the presence of conditioned medium from						
	Control	LCL _{HM}	LCL _{TL}	LCL _{D6}	LCL _{D10}	Raji	Molt-4f
Tonsil 1 Control	237	273	305	284	406	230	295
anti- μ	437	7,342	6,147	5,936	7,002	6,943	937
Tonsil 2 Control	520	684	495	406	537	394	466
anti- μ	1,034	12,360	10,436	9,825	7,894	9,062	2,146
Tonsil 3 Control	438	576	322	ND	ND	573	562
anti- μ	751	4,201	3,569	ND	ND	3,321	1,042

Tonsillar B-cells of high density were cultured as described for Fig. 1. F(ab')₂ anti- μ used throughout at 50 $\mu\text{g ml}^{-1}$. LCL_{HM} was studied shortly after establishment (1–3 months) and remained polyclonal over this period. LCL_{TL} is a long-term established cloned line of IgM λ specificity. LCL_{D6} and LCL_{D10} are medium term (~6 months) cloned lines of IgG κ specificity. Raji is an EBV-genome positive Burkitt's lymphoma line. Molt-4f was derived from a patient with T-cell leukaemia and retains T-cell properties. ND, not done.

cells may have been interfering with the assay (either by simple competition for the anti-immunoglobulin binding sites or through the formation of immune complexes *in situ*), we therefore used whole isotype-specific synergistic antibodies and looked for stimulating activity in supernatants from cloned lines of known immunoglobulin class. As shown in Table 1, B-cell stimulating activity could be readily demonstrated in LCL-CM in the absence of isotype interference. The magnitude of the invoked DNA synthesis was of the same order as that achieved with supernatants from mitogen-conditioned T cells. Tonsillar B lymphocytes of high density were unresponsive to all conditioned media unless anti-immunoglobulin was present. By contrast, cells of low density, already displaying appreciable DNA synthesis, were further activated with LCL-CM or conditioned T-cell supernatants in the absence of anti-immunoglobulin. These effects are similar to those described for physiological B-cell growth factor (BCGF) in human and murine systems^{2,4}.

To demonstrate reliably B-cell stimulating activity in conditioned medium from lymphoblastoid cells of any, or unknown, isotype we developed a delayed anti-immunoglobulin assay using a F(ab')₂ preparation of goat IgG specific for human μ

chains (Cappel) as co-stimulator antibody. This reagent was less directly mitogenic than the immobilized whole antibodies. In these experiments, the addition of conditioned media was delayed for 18 h following incubation of high density B cells with anti-immunoglobulin to avoid potential interference from supernatant immunoglobulins. Using this approach we demonstrated B-cell stimulating activity in LCL-CM over a range of anti- μ concentrations (Fig. 1). Again, the magnitude and distribution of the response of antibody-primed cells to LCL-CM followed closely that obtained when using T-cell supernatants as the source of stimulatory lymphokines. While individual tonsils responded variably, all B-cell preparations studied in the co-stimulator assay showed an enhanced DNA synthesis to LCL-CM over anti- μ alone (Table 2). Furthermore, all EBV-transformed lymphoblastoid lines tested released the stimulating activity (Table 2 and further data not shown). Raji, the one Burkitt's lymphoma line studied so far, also produced this activity. Conditioned medium collected from Molt-4f, a T-cell derived line included as a control in these studies, was slightly stimulatory for anti- μ primed tonsillar B lymphocytes although the magnitude of the response never approached the level achieved with supernatants from B-cell lines (Table 2).

We, and others, have recently identified an autocrine loop which sustains the proliferation of EBV-transformed cells in culture^{5,6}. Its maintenance is dependent on transferrin and an autogenous soluble activity supplied by cells at high density⁵. We considered therefore whether the perpetual growth of virally-transformed B lymphoblasts was a consequence of their utilizing their own BCGF and we have found that BCGF-containing T-cell supernatants could indeed replace the requirement for high cell density in LCL growth (experiments not detailed). More direct evidence was obtained during our attempts to isolate these activities from LCL conditioned medium by gel filtration chromatography. As shown in Fig. 2, both the activity which permitted S-phase entry of anti- μ primed tonsillar B cells (BCGF) and that which supported the growth of B lymphoblasts below their critical density (autostimulatory) eluted from a G-100 column as a single, relatively broad peak with an apparent molecular weight of 25,000–30,000. It is highly likely, then, that the two B-cell stimulating activities found in LCL-CM reside in the same molecular species. In addition, a similar elution profile on gel filtration has been reported for T-cell derived BCGF⁷. As conditioned media from this source contain conventional B-cell stimulating activity and also, as reported here, provide growth support for LCL, both activities from both cell types may be due to a common lymphocyte-derived growth factor normally produced only by T cells. EBV-transformed lymphoblastoid cell lines may provide a useful monoclonal source of BCGF.

We believe this to be the first report of a B-cell growth factor from cells of B-lymphocytic origin, but we concede that its induction may be a direct consequence of viral infection. Indeed, the constitutive production of an autostimulatory BCGF may be a crucial step in the events leading to the immortalization of these cells. An aberrant, or inappropriate, expression of proteins involved in cellular growth appears to be a regular feature of

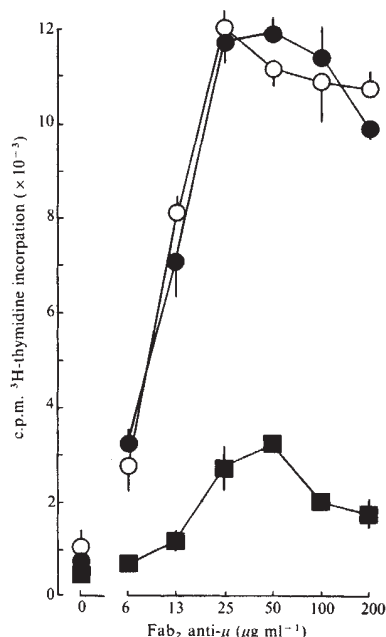


Fig. 1 Influence of LCL-CM and conditioned T-cell supernatants in a delayed co-stimulator assay. ■, Control medium; ●, LCL-CM_{HM}; ○, T-cell supernatants. High density tonsillar B cells were cultured at 10^5 for 18 h in microwells containing indicated concentrations of goat F(ab')₂ antibody specific for human μ chains (Cappel) in a final volume of 100 μl . On the addition of 50 μl conditioned medium, cells were returned to 37 °C for a further 56 h. Cells were pulsed with 0.5 μCi ^3H -thymidine for the final 16 h and the results are shown as means and standard deviations of quadruplicate determinations.

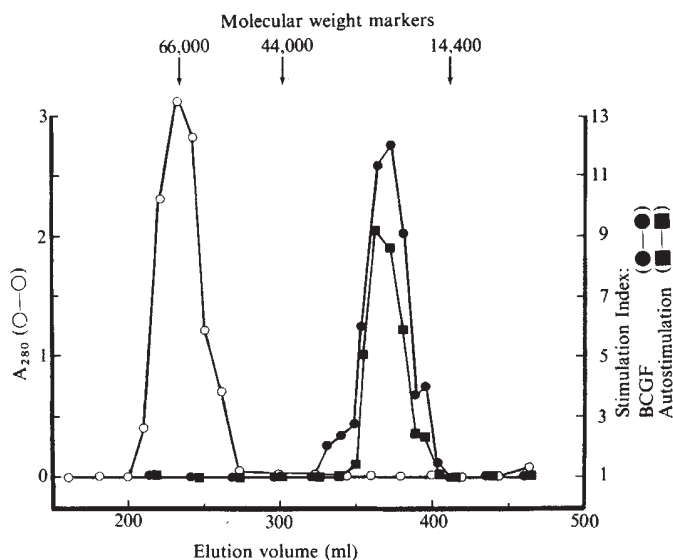


Fig. 2 Isolation of stimulating activities by gel filtration of conditioned medium. Conditioned medium was collected from LCL_{RM} seeded at 2×10^5 per ml and held at 37°C for 72 h in serum-free conditions—Iscove's modified Dulbecco's medium supplemented with bovine serum albumin (1 mg ml^{-1}) and transferrin ($25 \mu\text{g ml}^{-1}$; Gibco). Crude activity was precipitated from 75 ml of filtered supernatant by 90% ammonium sulphate at 4°C in the presence of 0.1 mg ml^{-1} of polyethylene glycol (PEG)₆₀₀₀. The precipitate was recovered by high speed centrifugation, reconstituted in 6 ml phosphate-buffered saline (PBS)/PEG (0.1 mg ml^{-1}) and dialysed against 400 ml PBS/PEG using Spectropor dialysis tubing (molecular weight cut-off 6,000). The sample was applied to a Sephadex G-100 column ($90 \text{ cm} \times 2 \text{ cm}$) calibrated with bovine serum albumin (66,000), ovalbumin (44,000) and lysozyme (14,000) and 7-ml fractions collected using PBS/PEG as the eluting buffer at a flow rate of 22 ml per h. In initial experiments, fractions corresponding to a particular molecular weight range from 70,000 to 6,000 were pooled and concentrated by pressure dialysis through PM10 (60,000–20,000 molecular weight) and YM2 membranes for fractions of molecular weight $< 20,000$. The concentrated fractions were finally dialysed against 50 volumes of RPMI 1640 containing 0.1 mg ml^{-1} PEG before being assayed for activity. Once it was established where the activities broadly resided, fractions within that area were collected and assayed directly without being pooled and concentrated (represented by elution volumes 320–414 ml). BCGF activity was assayed as described for Table 2 but with the fractions present at 33% from the start of culture. The autostimulatory activity was determined against LCL_{RM} cells plated in $150 \mu\text{l}$ of full medium at 10^3 per well (growth area 0.32 cm^2) in the presence or absence of collected fractions at 33% for 72 h. The stimulation index represents the augmentation of ^3H -thymidine uptake due to the presence of the indicated fractions. For the experiment shown, the background uptake for the BCGF assay (anti- μ primed tonsillar cells) was 730 c.p.m. and for the autostimulation assay, 474 c.p.m.

immortalized cells and a likely prerequisite of, although not sufficient to account for, the cancerous state^{8–14}. Further phenotypic alteration, providing some escape from normal host control, may be a vital additional element in the multi-stage evolution to full tumorigenicity^{15,16}. As Burkitt's lymphoma of endemic regions is closely associated with both the EBV-genome and an apparently deregulated oncogene (*c-myc*)^{3,17–19}, the finding that virally-transformed normal lymphocytes release a BCGF which is also autostimulatory, reveals some valuable insights into these areas and provides a functional basis for the role of EBV in the malignant evolution of this tumour.

We thank Dr G. Klein and the Department of Tumor Biology, Stockholm, for the supply of mycoplasma-free B95-8 virus supernatants, some of the cell lines used in this study and continuing encouragement. The Blood Group Reference Laboratory, Oxford, is acknowledged for financial support.

Received 14 March; accepted 25 May 1984.

- Andersson, J. & Melchers, F. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2479–2501 (1981).
- Howard, M. & Paul, W. E. *A. Rev. Immun.* **1**, 307–333 (1983).
- Nilsson, K. & Klein, G. *Adv. Cancer Res.* **37**, 319–380 (1982).
- Maruguchi, A., Butler, J. L., Kehrl, J. H. & Fauci, A. S. *J. exp. Med.* **157**, 530–546 (1983).
- Gordon, J., Ley, S. C., Melamed, M. D., Aman, P. & Hughes-Jones, N. C. *J. exp. Med.* **159**, 1554–1559 (1984).
- Blazar, B. A., Sutton, L. M. & Strome, M. *Cancer Res.* **43**, 4562–4569 (1983).
- Maizel, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 5998–6002 (1982).
- Uittenbogaart, C. H. & Fahey, J. L. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7004–7008 (1982).
- Twardzik, D. R. *et al. Science* **216**, 894–897 (1982).
- Kaplan, P. L., Anderson, M. & Ozanne, B. *Proc. natn. Acad. Sci. U.S.A.* **79**, 485–489 (1982).
- Anzano, M. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 6204–6268 (1983).
- Robbins, K. C. *et al. Nature* **305**, 605–608 (1983).
- Waterfield, M. C. *et al. Nature* **304**, 35–39 (1983).
- Downward, J. *et al. Nature* **307**, 521–527 (1984).
- Brickell, P. M. *et al. Nature* **306**, 756–760 (1983).
- Bernards, R. *et al. Nature* **305**, 776–779 (1983).
- Klein, G. *Cell* **32**, 311–315 (1983).
- Perry, R. P. *Cell* **33**, 647–649 (1983).
- Rabbits, T. H., Hamlyn, P. H. & Baer, R. *Nature* **306**, 760–765 (1983).
- Gordon, J., Melamed, M. D., Ley, S. C. & Hughes-Jones, N. C. *Immunology* **52**, 79–85 (1984).

Epidermal growth factor stimulates guanine nucleotide binding activity and phosphorylation of *ras* oncogene proteins

Tohru Kamata & James R. Feramisco

Cold Spring Harbor Laboratory, PO Box 100, Cold Spring Harbor, New York 11724, USA

Several human tumour cell lines contain genes that can transform NIH 3T3 cells into malignant cells. Certain genes have been classified as members of the *ras* oncogene family, namely, *Ha-ras*, *Ki-ras* or *N-ras*^{1–3}. The proteins encoded by the *ras* family are generally small (*Ha-ras*, for example, encodes a protein of molecular weight 21,000 named p21), and are associated with the inner surface of the plasma membrane. The only known biochemical property common to all forms of the *ras* proteins is the ability to bind guanine nucleotides, a property which may be closely related to the transforming ability of *ras* proteins^{4,5}. A GTP-dependent, apparent autophosphorylation (on threonine 59) activity has been identified only in the case of the v-*Ha-ras* protein⁶. Although the role of these biochemical activities in the transformation process remains unclear, we have initiated studies to determine the possible biochemical interactions of *ras* proteins with other membrane components. We report here the evidence that epidermal growth factor enhances the guanine nucleotide binding activity of activated c-*Ha-ras* or v-*Ha-ras* p21, and phosphorylation of v-*Ha-ras* p21, suggesting that some mitogenic growth factors may regulate those activities.

Membranes were isolated from a non-producer line of rat kidney cells (*Ha-NRK*) transformed by Harvey murine sarcoma virus (*Ha-MSV*) and preincubated with or without epidermal growth factor (EGF) followed by the addition of ^{32}P -labelled Mg-GTP^{2–}. Although no major changes in the total phosphoproteins in the membranes can be easily seen (Fig. 1a), immunoprecipitation of v-*Ha-ras* p21 showed that the phosphorylation of this protein was stimulated three- to fivefold by the addition of EGF (Fig. 1b). The recovery of v-*Ha-ras* p21 over the course of the reaction period as seen by the levels of ^{35}S -Met-labelled v-*Ha-ras* p21 determined by immunoprecipitation was the same in the presence or absence of EGF (Fig. 1c). Of importance in this reaction was the use of membranes isolated from serum-deprived *Ha-NRK* cells (24–48 h deprivation) which seemed to lower the basal (–EGF) phosphorylation state of the v-*Ha-ras* protein. If membranes from cells growing in medium containing 10% calf serum were used, little or no stimulation of phosphorylation of v-*Ha-ras* protein was observed whereas the basal level of phosphorylation was high. This probably reflects the presence of EGF (and perhaps other growth factors) in the serum. The phosphorylation enhancement of v-*Ha-ras* p21 due to EGF was dependent on the EGF concentration (Fig. 2a).

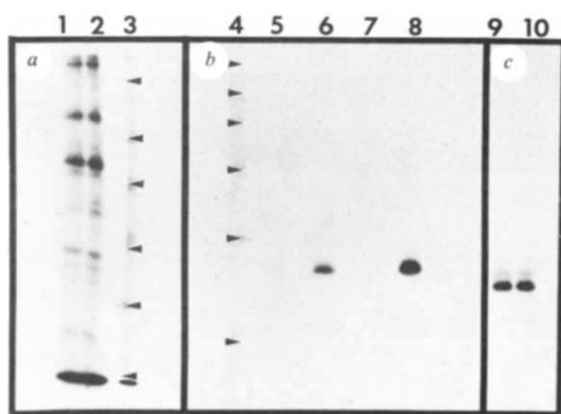
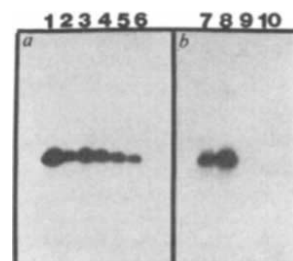


Fig. 1 Stimulation of phosphorylation of v-Ha-ras p21 in the presence of EGF. *a*, Total phosphoproteins detectable in the particulate fraction of HaNRK cells in the absence (lane 1) or presence (lane 2) of EGF analysed by a 10% SDS polyacrylamide gel. Molecular mass markers (lane 3) from top to bottom are: 200,000; 94,000; 68,000; 40,000; 30,000; and 14,000. *b*, Immunoprecipitated phosphorylated v-Ha-ras proteins immunoprecipitated from membranes treated without (lane 6) or with (lane 8) EGF. Lanes 5, 7, immunoprecipitations of the reaction mixtures without and with EGF, respectively, using a non-immune antibody. Molecular mass markers on the 12.5% SDS polyacrylamide gel are as indicated in *a*. *c*, 35 S-methionine-labelled v-Ha-ras proteins immunoprecipitated from reaction mixtures treated without (lane 9) or with (lane 10) EGF. **Methods:** HaNRK cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% calf serum. When the cells were confluent, they were rinsed twice in DMEM and maintained in DMEM (serum free) for 24–48 h. The cells were collected and washed twice with phosphate-buffered saline (PBS), then homogenized (dounce homogenizer) in 10 volumes of buffer A (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 0.1 mM EDTA, 0.1 mM 2-mercaptoethanol, 1 mM phenylmethylsulphonyl fluoride) and 10 volumes deionized H_2O . After pelleting the nuclei by low speed centrifugation, the supernatant was centrifuged at 100,000g for 60 min at 4°C. The particulate fraction was solubilized in buffer A containing 1% Triton X-100. Aliquots of the solubilized membranes (2.5×10^6 cell equivalent) were incubated with EGF (500 ng ml^{-1} in 1 mg ml^{-1} bovine serum albumin (BSA; BRL)) or BSA alone for 10 min at 4°C. The phosphorylation reaction was initiated by the addition of 10 μ Ci of [γ - 32 P]GTP (2,500 Ci $mmol^{-1}$, NEN) and $MgCl_2$ to 5 mM. The reaction was carried out for 10 min at 30°C. The reaction mixtures were preadsorbed with rabbit anti-rat IgG (2 μ l) and protein A-Sepharose (40 μ l of 50% suspension). The v-Ha-ras protein was immunoprecipitated by the addition of a rat anti-p21 monoclonal antibody (no. 259) (4 μ l). For control precipitations, a mouse monoclonal antibody against rat IgG (Fc) was used. (IgG isolated from a non-immunized rat also was used as a control and gave similar results.) After incubation for 1.5 h at 4°C, 4 μ l of rabbit anti-rat IgG and 40 μ l of a 50% suspension of protein A-Sepharose were added (1 h). The immunoprecipitates were washed with RIPA (–SDS) buffer (PBS, 1% sodium deoxycholate, 1% Triton X-100), boiled in Laemmli sample buffer and analysed by SDS-polyacrylamide gel electrophoresis (PAGE). Alternatively, HaNRK cells were labelled with 35 S-methionine (80 μ Ci ml^{-1} , NEN) for 18 h and the membranes were isolated as described above. After the membranes had been treated with EGF or BSA as above in the presence of 5 mM $MgCl_2$ and GTP (100 μ M) for 10 min at 4°C, v-Ha-ras p21 was immunoprecipitated by monoclonal antibody 259 as described above and analysed by SDS-PAGE followed by fluorography¹⁷.

HaNRK cell membranes treated with or without EGF were incubated with [γ - 32 P]ATP or [γ - 32 P]GTP to examine the nucleotide substrate specificity of the phosphorylation reaction. While the stimulatory effect of EGF on phosphorylation of v-Ha-ras p21 was detected in the presence of [γ - 32 P]GTP, no phosphorylated v-Ha-ras p21 was detected when [γ - 32 P]ATP was used as phosphoryl donor in the presence or absence of EGF (Fig. 2b). It seems that the EGF-induced enhancement of v-Ha-ras p21 phosphorylation is due either to a GTP-dependent

Fig. 2 Effect of EGF concentration on the phosphorylation of v-Ha-ras p21 and comparison of phosphoryl donors. *a*, Immunoprecipitated phosphorylated v-Ha-ras p21 from membranes treated with various concentrations of EGF: lanes 1 to 6 show the results using 500, 200, 100, 50, 25, 0 ng ml^{-1} of EGF, respectively. *b*, Immunoprecipitated phosphorylated v-Ha-ras p21 from membranes treated without (lane 7) or with (lane 8) EGF in the presence of [γ - 32 P]GTP. Lanes 9 and 10 are immunoprecipitations of the reaction mixtures treated without and with EGF, respectively, in the presence of [γ - 32 P]ATP.



Methods: HaNRK cell membranes were isolated and solubilized as described in Fig. 1 legend. After preincubation with varying concentrations of EGF (0–500 ng ml^{-1}), phosphorylation reactions were carried out in the presence of 10 μ Ci of [γ - 32 P]GTP and 5 mM $MgCl_2$ (*a*). Alternatively, after preincubation with 500 ng ml^{-1} of EGF, phosphorylation reactions were performed in the presence of 10 μ Ci of [γ - 32 P]GTP or [γ - 32 P]ATP (2,500 Ci $mmol^{-1}$, NEN) and 5 mM $MgCl_2$ for 10 min at 30°C (*b*). p21 was immunoprecipitated by anti-p21 antibody (no. 259) as described in Fig. 1 and analysed by SDS-PAGE.

protein kinase activity(s) activated by EGF or to a phosphatase activity inhibited by EGF. Although the stoichiometry and the site(s) of the phosphorylation of the v-Ha-ras protein are unknown, the chemical nature of the reaction has been determined. The immunoprecipitated phosphoproteins obtained from reactions done in the presence or absence of added EGF (as in Fig. 1) were treated with 1 M KOH for 2 h at 55°C⁷, and found to be completely labile, thus indicating a lack of phosphotyrosine. In addition, partial acid hydrolysis of the protein followed by two-dimensional phosphoamino acid analysis indicated the presence of phosphothreonine and phosphoserine, but not of phosphotyrosine (data not shown).

Other experiments were carried out to determine whether or not this effect occurs *in vivo*. In one such experiment, serum-deprived HaNRK cells (as described in Fig. 1) were incubated *in vivo* with 32 P-orthophosphate with or without 500 ng ml^{-1} of EGF. Immunoprecipitation of v-Ha-ras p21 showed a stimulation (about twofold) of phosphorylation of the protein in response to EGF (Fig. 3), a result similar to *in vitro* experimental results. Again, this increase in protein phosphorylation may well reflect the activation of a kinase or inhibition of a phosphatase by EGF, as EGF does not seem to alter the specific activity or pool size of (at least) 32 P-ATP in cells⁸.

One of the biochemical properties common to all *ras* proteins studied to date is guanine nucleotide binding activity⁴. This property can be detected either by immunoprecipitation assays or by direct measurements with purified proteins⁶. We examined the GDP binding activity of immunoprecipitated v-Ha-ras and the human activated c-Ha-ras protein, using isolated membrane fractions from HaNRK cells and Q4 cells, respectively, and the effect of EGF on this activity. Q4 is a rat cell line co-transfected with the adenovirus E1A gene and the human c-Ha-ras (T24) oncogene⁹. (Although the biochemical basis for the requirement for the adenovirus E1A gene (or one of several other genes) as well as the c-Ha-ras oncogene to transform primary rat cells is not understood^{9,10}, it seems that expression of the E1A gene alone does not interfere with EGF binding or stimulates the production of transforming growth factors by (at least) KB cells¹¹, which appears to be the only such cell line tested in these properties to date.) As shown in Fig. 4a, EGF doubles the 3H -GDP binding activity of the *ras* proteins present in both the HaNRK and Q4 membrane fractions. The stimulation depended on the EGF concentration and was complete at ~ 150 ng ml^{-1} (Fig. 4b). Attempts to determine if the guanine nucleotide binding activity of the *ras* proteins in human and rodent cell lines was also stimulated by EGF were unsuccessful due to the extremely low binding activity present in normal cell mem-

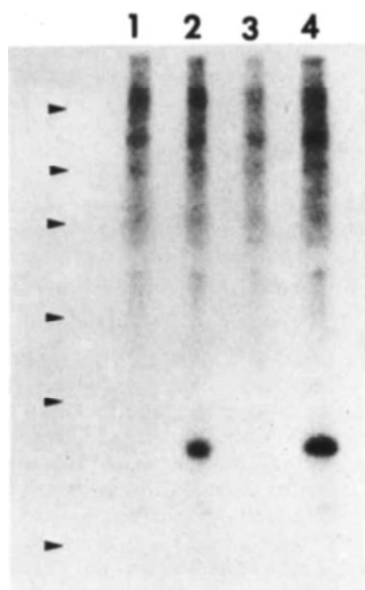


Fig. 3 Effect of EGF on v-Ha-ras p21 phosphorylation in HaNRK cells. Lanes 1, 2, immunoprecipitates from cells not treated with EGF; lanes 3, 4, immunoprecipitates from EGF-treated cells. Anti-p21 antibody was used in lanes 2 and 4 while a control antibody was used in lanes 1 and 3. Total incorporation of ^{32}P into the cell lysates used in lanes 1–4 was (in c.p.m. precipitated by trichloroacetic acid): 7.0×10^6 , 5.2×10^6 , 4.2×10^6 and 4.7×10^6 , respectively. Molecular weight markers are the same as indicated in Fig. 1.

Methods: After serum deprivation as described in Fig. 1, HaNRK cells (2.5×10^6) were re-fed with DMEM (serum free) and labelled with ^{32}P -orthophosphate (1 mCi ml^{-1}) for 90 min with or without EGF (500 ng ml^{-1}). The cells were lysed in RIPA (–SDS) buffer and the lysates were centrifuged at $100,000g$ for 60 min at 4°C . The supernatants were immunoprecipitated with $4 \mu\text{l}$ of anti-p21 monoclonal antibody (no. 259) or non-immune antibody as described in Fig. 1. The immunoprecipitates were analysed by SDS-PAGE.

branes, which made such measurements unreliable.

We have shown here that serum deprivation of v-Ha-ras-transformed cells (HaNRK) renders the *ras* protein in the membranes sensitive to EGF for the apparent *ras* autophosphorylation reaction. In addition, serum deprivation of cells transformed by either the activated human c-Ha-ras or murine v-Ha-ras proteins (Q4 or HaNRK) produces in *ras* proteins a sensitivity to EGF for guanine nucleotide binding activity. These results seem to indicate some type of biochemical interaction (either direct or indirect) between the EGF receptor system and the *ras* proteins. With the as yet sketchy but intriguing putative functional¹⁸ and structural¹⁹ similarities between the *ras* proteins and the G-like regulatory proteins involved in certain receptor systems, it will be useful to determine if the effects of EGF on the biochemical properties of the *ras* proteins have functional consequences in the cell. *Ras* transformed cells are known to have a reduced number of available binding sites for EGF compared with their normal cell counterparts¹². We have confirmed this for the HaNRK and Q4 cell lines used in this study (unpublished results). The apparent loss of EGF binding activity has been postulated to occur as a result of occupancy of the receptors by the α -type transforming growth factor (TGF_α), which is produced by the transformed cells¹³. If this is so, it will be of interest to see whether TGF_α , like EGF, alters the biochemical activities of the *ras* proteins through its binding to EGF receptor.

In an attempt to determine whether other growth hormones have similar effects on the biochemical activities associated with the *ras* proteins to those detected for EGF, we tested both insulin and platelet-derived growth factor (PDGF). Interestingly, in preliminary experiments, we found that insulin activates both

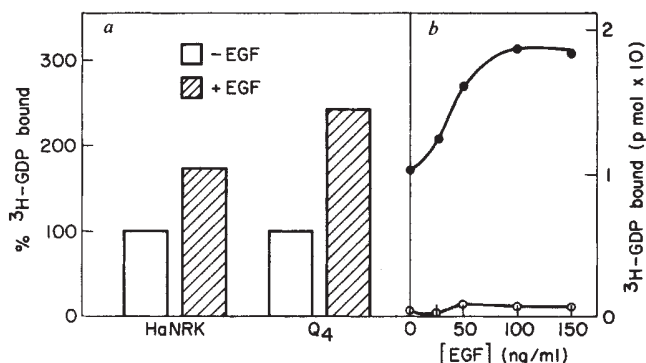


Fig. 4 Stimulation of ^3H -GDP binding activity of p21 by EGF. **a**, Immunoprecipitation of ^3H -GDP bound to p21 from HaNRK or Q4 cell membranes. The 100% value was 0.15 pmol (990 c.p.m.) for HaNRK and 0.068 pmol (449 c.p.m.) for Q4 cells. **b**, Immunoprecipitable ^3H -GDP bound to p21 from membranes treated with various concentrations of EGF, using anti-p21 antibody (no. 259) (●) and non-immune antibody (a mouse monoclonal antibody to rat IgG Fc) (○). Basal ^3H -GDP binding activity of p21 was 0.116 pmol (764 c.p.m.) and the radioactivity precipitated by non-immune antibody was $<0.01 \text{ pmol}$ (69 c.p.m.) and did not change significantly with the addition of EGF. Similarly, IgG isolated from preimmune rat serum gave no significant ^3H -GDP binding activity. Data show the mean value of duplicate experiments.

Methods: Membranes were isolated from serum-starved HaNRK or Q4 cells ($\sim 2.5 \times 10^6$ cells per data point) as described in Fig. 1 and preincubated with EGF (500 ng ml^{-1}) or BSA alone. $3 \mu\text{Ci}$ of ^3H -GDP (10 Ci mmol^{-1} , NEN), MgCl_2 to 5 mM final concentrations, and Triton X-100 to 1% final concentration were added (10 min at 4°C). ^3H -GDP bound to p21 was immunoprecipitated by anti-p21 antibody (no. 259) as described in Fig. 1 legend⁶, which includes 2.5 h further incubation at 4°C . The precipitates were washed four times with buffer A containing 5 mM MgCl_2 and analysed by liquid scintillation counting. Alternatively, HaNRK cell membranes (2.5×10^6 cells per data point) prepared as above were preincubated with various concentrations of EGF (0 – 150 ng ml^{-1}). ^3H -GDP binding assay was carried out in the presence of ^3H -GDP ($3 \mu\text{Ci}$) and 5 mM MgCl_2 as described above.

the phosphorylation state and guanine nucleotide binding property of v-Ha-ras p21 associated with HaNRK cell membranes, whereas PDGF has no effect (unpublished results). Studies to characterize the apparent insulin effect further are under way.

It seems that growth factors alter the state of tyrosine phosphorylation of at least two other membrane-associated oncogene or proto-oncogene proteins (EGF stimulated polyoma middle T antigen phosphorylation¹⁴, and PDGF apparently stimulates pp60^{c-src} phosphorylation; R. Ralston, personal communication). Considering the recent findings that the avian erythroblastosis virus oncogene protein (*erb-B*) may be a truncated version of the EGF receptor¹⁵ and that the simian sarcoma virus oncogene protein (*v-sis*) is related to PDGF¹⁶, it may be of interest to investigate interactions among various growth factors/receptor systems and the membrane-associated oncogene or proto-oncogene proteins.

We thank J. D. Watson for his support of this work; Dr Mike Wigler for his gift of cell lines and helpful discussions; Dr Bill Welch for helpful discussions; Drs Mark Furth and Ed Scolnick for their gift of monoclonal antibodies to p21; and Dr Earl Rulley for his gift of rat transformed cell lines. This work was supported by NIH grant GM2877 and NCI Cancer Center grant CA13106.

Note added in proof: Finkel and Cooper have reported the co-immunoprecipitation of the *ras* proteins and the transferrin receptor in human transformed cell lines²⁰.

Received 5 December 1983; accepted 14 May 1984.

- Der, C. J., Krontiris, T. G. & Cooper, G. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3637–3640 (1982).
- Parada, L. F., Tabin, C. J., Shih, C. & Weinberg, R. A. *Nature* **297**, 474–478 (1982).
- Shimizu, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 2112–2116 (1983).

4. Papageorge, A., Lowy, D. & Scolnick, E. *J. Virol.* **44**, 509–519 (1982).
5. Scolnick, E. M., Papageorge, A. G. & Shih, T. Y. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5355–5359 (1979).
6. Shih, T. Y., Papageorge, A. G., Stokes, P. E., Weeks, M. O. & Scolnick, E. M. *Nature* **287**, 686–691 (1980).
7. Eckhart, W., Hutchinson, M. A. & Hunter, T. *Cell* **18**, 925–933 (1979).
8. Chambard, J.-C., Franchi, A., LeCam, A. & Pouyssegur, J. *J. biol. Chem.* **258**, 1706–1713 (1983).
9. Ruley, H. E. *Nature* **304**, 602–606 (1983).
10. Land, H., Parada, L. F. & Weinberg, R. A. *Nature* **304**, 596–602 (1983).
11. Fisher, P. B., Boersig, M. R., Graham, G. M. & Weinstein, I. B. *J. cell. Physiol.* **114**, 365–370 (1983).
12. Todaro, G. J., DeLarco, J. E. & Cohen, S. *Nature* **264**, 26–31 (1976).
13. Anzano, M. A., Roberts, A. B., Smith, J. M., Sporn, M. B. & DeLarco, J. E. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6264–6268 (1983).
14. Segawa, K. & Ito, Y. *Nature* **304**, 742–744 (1983).
15. Downward, J. et al. *Nature* **307**, 521–527 (1984).
16. Waterfield, M. D. et al. *Nature* **304**, 35–39 (1983).
17. Bonner, W. M. & Laskey, R. A. *Eur. J. Biochem.* **46**, 83–88 (1974).
18. Kamata, T. K. & Feramisco, J. R. in *Cancer Cells* Vol. 1 (eds Levine, A., Toppe, W., Van de Woude, G. & Watson, J. D.) 11–16 (Cold Spring Harbor Laboratory, New York, 1984).
19. Gilman, A. G. *Cell* **36**, 577–579 (1984).
20. Finkel, T. & Cooper, G. M. *Cell* **36**, 1115–1121 (1984).

Role of soluble myosin in cortical contractions of *Xenopus* eggs

Kathy Christensen, Roger Sauterer & R. W. Merriam*

Department of Neurobiology and Behavior, State University of New York, Stony Brook, New York 11794, USA

The layer of cytoplasm underlying the plasmalemma of *Xenopus* eggs has contractile activity which is of vital importance in fertilization and early development, being involved in such processes as sperm engulfment, cortical granule exocytosis, development of the axes of embryonic symmetry and cleavage^{1–3}. In amphibian eggs this layer is also involved in wound healing and changes of cellular shape at gastrulation^{4–6}. Two kinds of contractile structures can be distinguished near the surface of *Xenopus* eggs^{7,8}. To characterize the mechanism and regulation of this contractile activity, we have experimentally induced cortical contractions in bisected living *Xenopus* eggs. We have shown previously that cortical contractions are induced by calcium ions in the bisected egg⁷. Here we show that extraction of soluble cytoplasmic components prevents the calcium-induced contractions, but that addition of exogenous soluble myosin restores them. In oocytes, both soluble and insoluble components of the cortical cytoplasm are unable to support contraction. Thus, during meiotic maturation of oocytes into eggs, both of the components of the cortical cytoplasm must change so as to become competent for contraction.

When *Xenopus* eggs with the vitelline membrane removed are bisected in a contraction buffer containing micromolar calcium, surface contractions are generated in 1–4 min which cause a decrease in surface area in both animal and vegetal hemispheres (Fig. 1). When half-cells are extracted for 20 min in buffer containing EGTA and ATP but no calcium, neither induced nor

normal contractions occur and the subsequent addition of calcium causes little or no contraction (Table 1a). If, however, the half-cells are extracted for 20 min in buffer containing the soluble components of egg homogenates and ATP (see Table 1), subsequent addition of calcium induces contraction (Table 1a). Soluble components from the cytoplasm are thus necessary for the contraction.

Nucleotide requirements were explored by transdialysing small molecular weight metabolites out of the soluble fraction of egg homogenates; the dialysed extract could not then support contraction. When ATP or GTP was added to the dialysed extract, contraction again resulted; ADP did not support the contractions (Table 1b).

The half-cell contraction system was used as an assay for the isolation and characterization of the non-dialysable components of egg homogenates that are required for contraction. Using a combination of anion-exchange and gel-filtration chromatography, a protein was isolated which could support contraction in the presence of calcium and ATP (Table 1c). Analysis on SDS-polyacrylamide gels showed that the major polypeptide of this protein co-migrated with the heavy chain of rabbit muscle myosin (Fig. 2). Highly purified rabbit skeletal muscle myosin (Fig. 2) was then added to the extracted half-cells to determine whether it would support contraction; it did, as is shown in Table 1d.

To confirm that soluble myosin was indeed the required soluble factor, we used *N*-ethylmaleimide-modified heavy meromyosin (NEM-HMM) from rabbit skeletal muscle. This compound has been shown to bind irreversibly to actin microfilaments, blocking myosin-mediated contractions⁹. Table 1e shows that preincubation of half-cells in NEM-HMM caused a marked inhibition of calcium-induced contraction. We conclude that the soluble components necessary for calcium-induced contraction of the cortical cytoskeleton are calcium, ATP and soluble myosin. We are unable at present to define which of the two superficial contractile systems^{7,8} has the requirement for soluble myosin, or whether both do.

The myosin requirement is neither tissue- nor species-specific. To our knowledge this is the first time that an extracted cell model^{10–15} or isolated cell structure¹⁶ has been shown to require exogenous, soluble myosin for an induced contraction. One possible interpretation of these results is that this egg myosin is phosphorylated and hence free of the cytoskeleton¹⁶; if this is the case, the calcium trigger may act at a different level from light-chain phosphorylation.

Ripe ovarian oocytes are known to be refractory to activation and calcium-induced contraction^{17,18}. We were curious to know whether this contractile incompetence was due to the insoluble cytoskeleton, the soluble myosin or both. Experiments were undertaken to examine the effect of active egg myosin on the extracted oocyte cytoskeleton and the effect of soluble oocyte components on the active egg cytoskeleton. Results from several experiments on stage VI¹⁹ oocytes are summarized in Table 2. Clearly, neither oocyte cytoskeleton nor oocyte soluble components can support cortical contractions as induced by exogenous calcium in the half-cell system.

* To whom correspondence should be addressed.

Fig. 1 Cortical contractions, induced in *Xenopus* half-cells. *a*, Animal hemisphere before contraction. *b*, The same hemisphere after contraction. $\times 25$. Unfertilized eggs were dejellied in 35 mM mercaptoethanol, pH 8.9. The vitelline membranes were removed and eggs cut in half with fine forceps in a contraction buffer (25 mM Tris-HCl, pH 7.4, 80 mM NaCl, 2 mM MgCl₂, 3 mM CaCl₂, 2 mM EGTA, 0.05 mM ATP, 0.1 mM PMSF, 0.02% NaN₃).

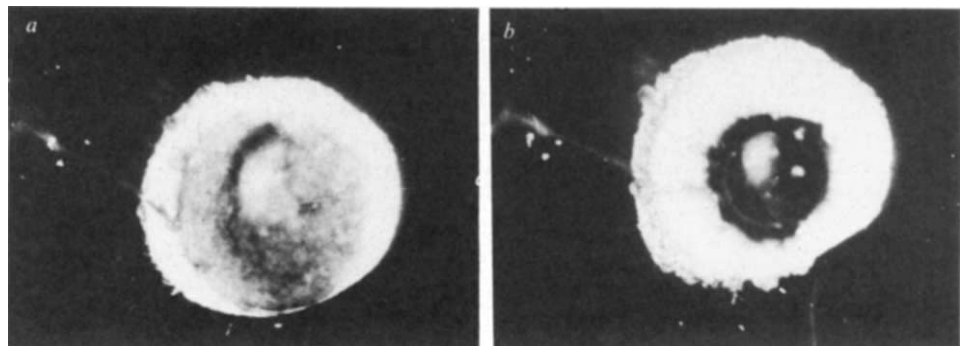


Table 1 Soluble factors required for contraction of the insoluble cytoskeleton in half-cells of *Xenopus*

	Half-cells extracted in:	Extraction time before calcium (min)	Additions to buffer	No. of cells tested	% Cells contracting
<i>a</i>	Buffer	20	—	42	10
	Buffer	20	ATP	15	0
	Egg extract	20	ATP	42	98
<i>b</i>	Dialysed egg extract	20	—	21	5
	Same	20	ATP	21	86
	Same	20	GTP	20	90
	Same	20	ADP	20	5
<i>c</i>	Isolated soluble egg factor	20	—	153	10
	Same	20	ATP	153	82
<i>d</i>	Purified skeletal muscle myosin	20	—	50	8
	Same	20	ATP	50	96
<i>e</i>	NEM-BSA	10	ATP	50	94
	NEM-HMM	10	ATP	80	18
	Egg myosin + NEM-BSA	20	ATP	30	87
	Egg myosin + NEM-HMM	20	ATP	60	15

a, Buffer: 25 mM Tris-HCl pH 7.4, 80 mM NaCl, 2 mM MgCl₂, 2 mM EGTA, 2 mM β -mercaptoethanol, 0.1 mM phenylmethylsulphonyl fluoride (PMSF), 0.02% NaN₃. Egg extract: Dejellied eggs were homogenized in 1.5 vol. of cold imidazole buffer (25 mM imidazole pH 6.4, 1 mM NaCl, 1 mM β -mercaptoethanol, 2 mM EGTA, 0.02% NaN₃). The homogenate was centrifuged for 50 min at 25,000g at 5 °C and the supernatant further centrifuged for 2 h at 100,000g at 2 °C. The supernatant is called the egg extract. *b*, Dialysed egg extract: Four volumes of imidazole buffer were transdialysed through egg extract, holding the volume constant. A membrane filter with a molecular weight cut-off of 1,000 was used. *c*, Isolated soluble egg factor: Egg extract was adsorbed on a short DEAE-cellulose column in imidazole buffer, pH 6.4. Adsorbed anions were eluted with 1.5 M NaCl and passed through a Sephadex G-150 column (bed volume 230 ml) equilibrated with buffer (*a*) containing 1.5 M NaCl. A peak eluting just after the void volume contained the soluble egg factor (see Fig. 2). *d*, Purified skeletal muscle myosin: Rabbit skeletal muscle myosin was prepared essentially by the procedure of Offer *et al.*²⁸ or Kielley and Harrington²⁹. For half-cell assays, the myosin was further purified by passing it through a Sephadex G-150 column equilibrated with buffer (*a*) containing 1.5 M NaCl. *e*, NEM-BSA, *N*-ethylmaleimide was coupled to bovine serum albumin to serve as a control; it was used at 3 mg ml⁻¹. NEM-HMM, heavy meromyosin was prepared from rabbit skeletal muscle myosin purified by the method of Offer *et al.*²⁸ and coupled with *N*-ethylmaleimide by the procedure of Meeusen *et al.*⁹. It was used at 3 mg ml⁻¹.

Table 2 Efficacy of cortical matrix and soluble factors from oocytes in calcium-induced contractions

Cortical matrix source	Extract source	No. of cells tested	% Cells contracting
Egg	Egg	13	100
Egg	Oocyte	45	2
Oocyte	Egg	35	0
Oocyte	Oocyte	12	0

Stage VI oocytes, used here, are the largest size of ovarian cells. Half-cells were extracted 18 min before contraction was induced by the addition of 3 mM CaCl₂ (to buffer containing 2 mM EGTA and 0.05 mM ATP).

The meiotic maturation of oocytes to metaphase II is accompanied by ovulation, stimulation of protein synthesis and a chain of events resulting in an activatable (and contractile) egg²⁰⁻²². Our results show that contractile competence in response to calcium also arises in the 6–14 h required for meiotic maturation. Because of the known control of non-muscle myosins through myosin phosphorylation^{23,24} and the involvement of phosphate exchanges during *Xenopus* oocyte meiotic maturation²⁵, it is tempting to speculate that protein phosphate exchanges have a role in the development of contractile competency on maturation of an oocyte to an egg. It is also possible that the calcium sequestering and releasing system comes into being during oocyte maturation^{26,27}.

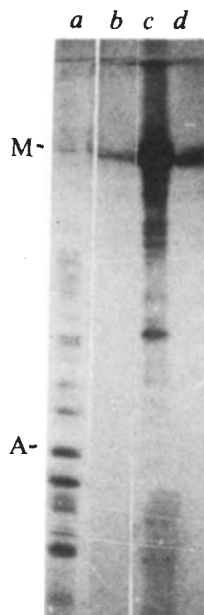


Fig. 2 SDS-polyacrylamide gel electrophoresis of egg and skeletal muscle myosin preparations. Coomassie blue stain. *a*, Total egg extract; *b*, purified egg soluble myosin after Sephadex G-150 chromatography; *c*, rabbit skeletal muscle myosin after ammonium sulphate precipitation²⁹; *d*, rabbit skeletal muscle myosin after final Sephadex purification. M indicates the position of the myosin heavy chain, A the position of actin.

Received 28 February; accepted 3 May 1984.

- Gerhart, J., Black, S., Gimlich, R. & Scharf, S. in *Time, Space and Pattern in Embryonic Development* (eds Jeffery, W. R. & Raff, R. A.) 261–286 (Liss, New York, 1983).
- Vacquier, V. D. *Dev Biol.* **84**, 1–26 (1981).
- Elinson, R. P. *Dev Biol.* **47**, 257–268 (1975).
- Holtfreter, J. *J. exp. Zool.* **93**, 251–323 (1943).
- Holtfreter, J. *J. exp. Zool.* **94**, 261–317 (1943).
- Merriam, R. W. & Christensen, K. *J. Embryol. exp. Morph.* **75**, 11–20 (1983).
- Merriam, R. W. & Sauterer, R. A. *J. Embryol. exp. Morph.* **76**, 51–65 (1983).
- Merriam, R. W., Sauterer, R. A. & Christensen, K. *Dev Biol.* **95**, 439–446 (1983).
- Meeusen, R. L. & Cande, W. Z. *J. Cell Biol.* **82**, 57–65 (1979).
- Hoffman-Berling, H. *Biochim. biophys. Acta* **14**, 182–194 (1954).
- Hoffman-Berling, H. *Biochim. biophys. Acta* **19**, 453–463 (1956).
- Holzappel, G., Wehland, J. & Weber, K. *Expl Cell Res.* **148**, 117–126 (1983).
- Cande, W. Z., Tooth, P. J. & Kendrick-Jones, J. *J. Cell Biol.* **97**, 1062–1071 (1983).
- Burnside, B., Smith, B., Nagata, M. & Porrello, K. *J. Cell Biol.* **92**, 199–206 (1982).
- Owaribe, K., Kodama, R. & Eguchi, G. *J. Cell Biol.* **90**, 507–514 (1981).
- Broschat, K. O., Stidwill, R. P. & Burgess, D. R. *Cell* **35**, 561–571 (1983).
- Moreau, M., Doree, M. & Guerrier, P. *J. exp. Zool.* **197**, 443–449 (1976).
- Belanger, A. M. & Schuetz, A. W. *Dev Biol.* **45**, 378–381 (1975).
- Dumont, J. N. *J. Morph.* **136**, 153–180 (1972).
- Masui, Y. & Clarke, H. J. *Int. Rev. Cytol.* **57**, 185–282 (1979).
- Wasserman, W. J., Richter, J. D. & Smith, L. D. *Dev Biol.* **89**, 152–158 (1982).
- Dettlaff, T. A. *J. Embryol. exp. Morph.* **16**, 183–195 (1966).
- Scholey, J. M., Taylor, K. A. & Kendrick-Jones, J. *Nature* **287**, 233–235 (1980).
- Kuczmarski, E. R. & Spudich, J. A. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7292–7296 (1980).
- Huchon, D., Ozon, R. & Demaille, J. G. *Nature* **294**, 358–359 (1981).
- Campanella, C. & Andreucetti, P. *Dev Biol.* **56**, 1–10 (1977).
- Gardiner, D. M. & Grey, R. D. *J. Cell Biol.* **96**, 1159–1163 (1983).
- Offer, G., Moos, C. & Starr, T. J. *molec. Biol.* **74**, 653–676 (1973).
- Kielley, W. W. & Harrington, W. F. *Biochim. biophys. Acta* **41**, 401–421 (1960).

Quantification of the close association between DNA haplotypes and specific β -thalassaemia mutations in Mediterraneans

Haig H. Kazazian Jr*, Stuart H. Orkin†, Alexander F. Markham‡, Catherine R. Chapman*, Hagop Youssoufian* & Pamela G. Waber*

* Department of Pediatrics, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA

† Division of Hematology and Oncology, Children's Hospital Medical Center and the Sidney Farber Cancer Institute, and the Department of Pediatrics, Harvard Medical School, Boston, Massachusetts 02115, USA

‡ ICI Pharmaceuticals Division, Mereside, Alderly Park, Macclesfield, Cheshire SK10 4TG, UK

It has been suggested that there is a close linkage between specific restriction fragment polymorphism patterns, defined as haplotypes, in the β -globin gene cluster and specific mutations in Mediterranean people with thalassaemia¹. This association formed the basis of a strategy for the efficient characterization of β -thalassaemia mutations from the DNA sequence of one or two β -thalassaemia genes derived from each haplotype in each ethnic group. Subsequently, Robertson and Hill argued that this strategy greatly underestimates the number of mutations on haplotypes which are frequent among normal chromosomes². We have therefore now analysed the proposed association and strategy quantitatively by the use of oligonucleotide hybridization and direct restriction analysis. Our results suggest that: (1) the association of specific haplotypes with specific mutations is high, but not invariant; (2) a different β -thalassaemia mutation has arisen within each haplotype in Mediterraneans; and (3) mutation spread from one haplotype to another occurs mainly through meiotic recombination within a 9-kilobase region 5' to the β -globin gene.

We have previously used restriction fragment (RF) polymorphisms in the β -gene cluster to define haplotypes in normal and β -thalassaemia chromosomes in various ethnic groups³. We found three common sequences for the 32 kilobases (kb) between the ϵ - and δ -globin genes and four common sequences for the 18 kb beginning at the 5' end of the β gene and extending

in the 3' direction^{3,4}. Between these two sequence clusters lies a region of 9 kb within which the recombination rate has been estimated to be 30 times the average recombination rate for a segment of this size (A. Chakravarti *et al.*, in preparation)³. Within the β gene itself three common DNA sequences, called frameworks, have been observed in Mediterraneans. Frameworks 1 and 2 differ by a single nucleotide (nt) substitution in intervening sequence 2 (IVS-2) of the β gene, while framework 3 has the substitution of framework 2 plus four additional ones. In Mediterraneans, haplotypes I, II and IX contain framework 1 β genes, haplotypes III, Va, Vb and VIII have framework 2, and haplotypes IV, VI, VII and X have framework 3 β genes¹ (see Table 1).

The initial suggestion of a close association between specific haplotypes in the β -gene cluster and specific β -thalassaemia mutations in Mediterraneans was based on sequence analysis of 13 β -thalassaemia genes and direct detection of the mutation by restriction endonuclease analysis in 15 others¹. The association of haplotype and mutation was then used to develop a strategy for the characterization of β -thalassaemia mutations in other ethnic groups⁵⁻⁷. Subsequently, Robertson and Hill argued that, since each β -thalassaemia mutation originated on the background of a normal haplotype, this strategy greatly underestimates the number of β -thalassaemia mutations on haplotypes which occur frequently in normal chromosomes². Because of the potential general applicability of our approach to other autosomal gene recessive disorders, we sought to determine the specific β -thalassaemia mutation present in each of 162 Mediterranean β -thalassaemia chromosomes available to us.

Oligonucleotides (19-mers) were chemically synthesized specific for each of the following mutations⁸: (1) IVS-1 nt 110 (G-A)⁹; (2) nonsense codon 39 (C-T)^{10,11}; (3) IVS-1 nt 1 (G-A)¹; and (4) IVS-1 nt 6 (T-C)¹. They were 5'-end labelled with ³²P by polynucleotide kinase and used to detect the presence or absence of each mutation in genomic DNA of 162 unselected Mediterraneans who were heterozygous for β -thalassaemia. Heterozygous individuals were studied so that a positive hybridization result could unambiguously assign a specific allele to each β -thalassaemia gene (Fig. 1). Most of the assigned alleles were then tested with one or more of the three remaining oligonucleotide probes and gave negative hybridization results. Known β -thalassaemia mutations IVS-2 nt 1 (ref. 12), IVS-2 nt 745 (ref. 1), -87 (ref. 1) and frameshift codon 6 (ref. 13) were detected directly by restriction endonuclease analysis with *Hph*I, *Rsa*I, *Avr*II and *Mst*II, respectively.

Table 1 Association of β -thalassaemia mutations and chromosomal backgrounds in Mediterraneans

β -gene framework	Haplotype	Mutations									Uncharacterized*
		IVS-1 nt 110	Nonsense 39	Frame-shift 6	IVS-2 nt 1	IVS-1 nt 1	-87	Frame-shift 8	IVS-1 nt 6	IVS-2 nt 745	
1	I	51	9	1							2
	II	1	32								
	IX	1	2	1							2
2	III				8						
	Va			1		15					2
	Vb				4						
	VIII						2				
3	IV							1			
	VI								15		
	VII		1						1	9	
	X								1		

Known, 156 (96%); total, 162. Haplotypes I to IX are those presented by Orkin *et al.*¹. Haplotypes Va and Vb differ at the polymorphic *Hinf*I site 5' to the β gene. Haplotype X is - + - - - + at the seven sites presented in Fig. 1 of that paper. The 13 β -thalassaemia genes sequenced by Orkin *et al.*¹ are included in the 156 characterized genes. One or more of the four uncharacterized framework 1 β genes (haplotypes I and IX) may contain the $\beta^{K_{\text{nosso}}}$ gene (codon 27 (G-T))¹⁶ which has previously been found associated with haplotype I (S.H.O. and S. E. Antonarakis, unpublished results) and has not been looked for in our panel. Another characterized allele, the IVS-2 nt 705 mutation¹⁷, is probably not among our 162 β -thalassaemia genes as it was originally described in a framework 3 β gene, and all 28 framework 3 β genes available to us have been characterized. The two uncharacterized framework 2 β genes (haplotype Va) presumably represent an as yet uncharacterized allele.

Table 2 Frequency of specific haplotypes in β^A and β^{thal} chromosomes

Haplotype	β^A chromosomes (%)	β^{thal} chromosomes (%)
I	42 (34)	63 (39)
II	8 (7)	33 (20)
III	8 (7)	8 (7)
IV	4 (3)	1 (1)
Va	22 (18)	18 (11)
Vb	—	4 (2)
VI	—	15 (9)
VII	12 (10)	11 (7)
VIII	—	2 (1)
IX	15 (12)	6 (4)
X	—	1 (1)
Others	11 (9)	—
Total	122	162

Using these methods we have determined the mutation in 156 of 162 (96%) β -thalassaemia genes of Mediterraneans available in our panel (Table 1). Of 156 genes identified, 9 mutations were initially correlated with RF haplotype after DNA sequence analysis of 14 β -globin genes^{1,13}. Knowledge of these nine mutations and their associated haplotypes was then used to determine the mutation type of 142 other β -thalassaemia genes, 22 by restriction analysis and 120 by hybridization with oligonucleotide probes. The remaining six genes are still uncharacterized, even though they have been studied for the common β -thalassaemia mutations found in this region.

We have previously shown that since the same β^A chromosome haplotypes are found in different racial and ethnic groups, they must have originated before the divergence of the human races³. In contrast, as each ethnic group afflicted with β -thalassaemia has its own battery of mutations, these disease-producing mutations have occurred since the divergence of the races^{1,5-7}. Comparison of haplotypes of β^A and β^{thal} chromosomes in Mediterraneans demonstrates their similarity (Table 2). Tables 1 and 2 show that the common β -thalassaemia mutations, IVS-1 nt 110 and nonsense codon 39, originated in common β^A chromosome backgrounds (haplotypes I and II). These two mutations were probably the first to experience positive selection pressure in this region.

Taken as a whole, these data demonstrate a high association of specific haplotypes with specific mutations. On average, 86% of the mutations within a particular haplotype are identical, while 86% of the occurrences of a particular mutation are associated with a particular haplotype. Thus, the simplest interpretation of these data is that a different β -thalassaemia mutation has arisen within each normal chromosome background, haplotypes I to IX. For example, the IVS-1 nt 110 mutation occurred in haplotype I, IVS-1 nt 1 in haplotype Va, and so on. Five of the nine mutations (IVS-1 nt 110, nonsense codon 39, frameshift codon 6, IVS-2 nt 1 and IVS-1 nt 6) have spread to other chromosome backgrounds, probably through meiotic recombination (see below). Mutation spread has almost never involved a change of β -gene framework. For example, the IVS-1 nt 110 mutation has spread from haplotype I to haplotypes II and IX, all of which contain β -gene framework 1.

Of 10 examples of mutation spread (Table 1), 8 can be accounted for by the mechanism of recombination between the polymorphic *TaqI* site 5' to the δ -globin gene and the polymorphic *HgiAI* site in the first exon of the β -globin gene (A. Chakravarti *et al.*, in preparation)^{1,3,14}. The two exceptions are the spread of the nonsense codon 39 mutation from haplotype I to haplotype VII (framework 1 to framework 3) and the frameshift codon 6 mutation from either haplotype I or IX to haplotype Va (framework 1 to framework 2). These examples of mutation spread from one β -gene framework to another within an ethnic group may be due to inter-allelic gene conversion events or recurrent mutation, as discussed elsewhere for β^S - and β^E -globin genes^{4,15}.

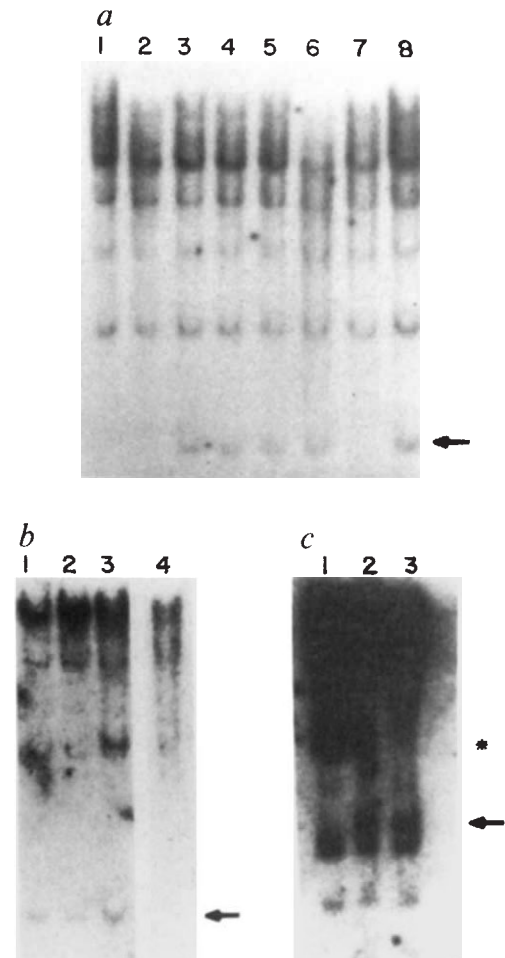


Fig. 1 Detection of three β -thalassaemia alleles by oligonucleotide hybridization. 10 μ g of genomic DNA of Mediterraneans heterozygous for β -thalassaemia were digested with *Bam*HI, subjected to electrophoresis in 1% agarose gels, and Southern transferred to nitrocellulose filters¹⁸. Hybridization with ³²P-oligonucleotides (19-mers) specific for various β -thalassaemia mutations and washing of filters were carried out as described previously^{19,20}. The mutations examined occur in the 5' one-third of the β -globin gene which is included in a 1.8-kb *Bam*HI fragment denoted by an arrow in *a*, *b* and *c*. In *a*, the probe is specific for the IVS-1 nt 6 mutation and positive hybridization is seen in lanes 3, 4, 5, 6 and 8. In *b*, the probe is specific for the IVS-1 nt 110 mutation and positive hybridization is present in lanes 1, 2 and 3. In *c*, the probe is specific for the IVS-1 nt 1 mutation and positive hybridization is present in lanes 2 and 3. Note the presence in *c* of a non-globin polymorphic fragment denoted by the asterisk.

Since only 6 of 162 β -thalassaemia genes in our panel remain outstanding, the number of alleles still uncharacterized in Mediterraneans is small, perhaps two or less. Alleles which occur with a significant frequency in this population have been discovered. It is noteworthy that this result was achieved after sequence analysis of only 14 β -thalassaemia genes^{1,13}, or less than 10% of those available to us, further demonstrating the power of the strategy for the molecular characterization of this disease.

The strong association between haplotype and mutation, such that for even the most common haplotypes almost all mutations are of a single kind, is somewhat surprising. The association is presumably due to (1) the recent onset of positive selection for β -thalassaemia heterozygotes in malaria-infested regions of the Mediterranean Basin and (2) a significant lapse of time between the occurrence of the first and any subsequent mutations on a particular chromosome background. On the other hand, a close association of specific haplotypes with specific mutations may

not pertain in single-gene disorders in which there is no positive selection for heterozygotes².

This work was supported by research grants from the NIH and the March of Dimes-Birth Defects Foundation to H.H.K. and S.H.O., and an RCDA to S.H.O. We thank Emily Pasterfield for help in manuscript preparation.

Received 5 January; accepted 30 March 1984.

- Orkin, S. H. *et al.* *Nature* **296**, 627-631 (1982).
- Robertson, A. & Hill, W. G. *Nature* **301**, 176-177 (1983).
- Antonarakis, S. E. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **79**, 137-141 (1982).
- Kazazian, H. H. Jr *et al.* *Am. J. hum. Genet.* **36**, 212-217 (1984).
- Antonarakis, S. E. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **81**, 1154-1158 (1984).
- Kazazian, H. H. Jr *et al.* *EMBO J.* (in the press).
- Cheng, T.-c. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **81**, 593-596 (1984).
- Wallace, R. B. *et al.* *Nucleic Acids Res.* **9**, 3647-3656 (1981).
- Spritz, R. A. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **78**, 2455-2459 (1981).
- Moschonas, N. *et al.* *Nucleic Acids Res.* **9**, 4391-4401 (1981).
- Trecartin, R. F. *et al.* *J. clin. Invest.* **68**, 1012-1017 (1981).
- Treisman, R. A. *et al.* *Cell* **29**, 903-911 (1982).
- Kazazian, H. H. Jr *et al.* *Am. J. hum. Genet.* **35**, 1028-1033 (1983).
- Maeda, N., Bliska, J. B. & Smithies, O. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5012-5016 (1983).
- Antonarakis, S. E. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **81**, 853-856 (1984).
- Arous, N. *et al.* *FEBS Lett.* **147**, 247-250 (1983).
- Dobkin, C. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **80**, 1184-1188 (1983).
- Southern, E. M. *et al.* *J. molec. Biol.* **98**, 503-517 (1975).
- Conner, B. J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **80**, 278-282 (1983).
- Orkin, S. H., Markham, A. F. & Kazazian, H. H. Jr *J. clin. Invest.* **71**, 775-779 (1983).

DNA sequences of telomeres maintained in yeast

Janis Shampay*, Jack W. Szostak† & Elizabeth H. Blackburn*‡

*Department of Molecular Biology, University of California, Berkeley, California 94720, USA

†Dana-Farber Cancer Institute, and Department of Biological Chemistry, Harvard Medical School, Boston, Massachusetts 02115, USA

Telomeres, the ends of eukaryotic chromosomes, have long been recognized as specialized structures. Their stability compared with broken ends of chromosomes^{1,2} suggested that they have properties which protect them from fusion, degradation or recombination^{1,3,4}. Furthermore, a linear DNA molecule such as that of a eukaryotic chromosome must have a structure at its ends which allows its complete replication⁵⁻⁸, as no known DNA polymerase can initiate synthesis without a primer. At the ends of the relatively short, multi-copy linear DNA molecules found naturally in the nuclei of several lower eukaryotes, there are simple tandemly repeated sequences⁹⁻¹⁷ with, in the cases analysed, a specific array of single-strand breaks, on both DNA strands, in the distal portion of the block of repeats^{9,10,17}. In general, however, direct analysis of chromosomal termini presents problems because of their very low abundance in nuclei. To circumvent this problem, we have previously cloned a chromosomal telomere of the yeast *Saccharomyces cerevisiae* on a linear DNA vector molecule¹⁸. Here we show that yeast chromosomal telomeres terminate in a DNA sequence consisting of tandem irregular repeats of the general form C₁₋₃A. The same repeat units are added to the ends of *Tetrahymena* telomeres, in an apparently non-template-directed manner, during their replication on linear plasmids in yeast. Such DNA addition may have a fundamental role in telomere replication.

Figure 1a shows a map of the linear yeast plasmid pSZ219. One end is a telomere from a yeast chromosome and the other end is derived from the terminal restriction fragment of the naturally occurring linear ribosomal DNA molecule from the somatic nucleus of the ciliate *Tetrahymena*¹⁸. To obtain amounts of yeast telomeric DNA sufficient for sequence analysis, the yeast telomere on pSZ219 was subcloned in the circular *Escherichia coli* vector pBR322 as described in Fig. 1 legend. The resulting recombinant molecules containing the yeast

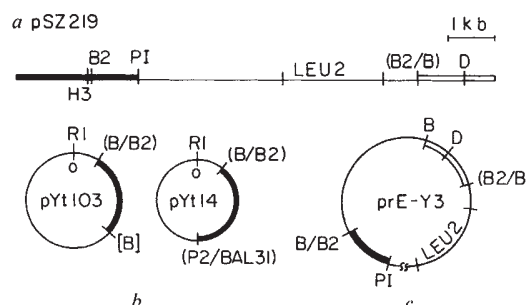


Fig. 1 Subcloning of telomeric fragments from a linear yeast plasmid. *a*, The linear plasmid pSZ219-5 (ref. 18) consists of a yeast telomere (solid bar on left) and the terminal fragment of the linear ribosomal RNA gene (rDNA) of *Tetrahymena* (open bar on right). A selectable marker (*LEU2*) is present, with pBR322 sequences, in the central portion of the molecule (thin line). *b*, *c*, Recombinant plasmids carrying the yeast telomeric fragment (*b*) or the *Tetrahymena* rDNA-derived telomere (*c*). pSZ219-5 was purified from log-phase yeast cells by a procedure modified from ref. 29, and digested with *S*₁ nuclease in conditions recommended by the manufacturer (BRL) to remove the extreme molecular termini^{9,11,17}. *Bam*HI linkers (New England Biolabs) were ligated onto the resulting blunt ends, and the pSZ219 DNA was then digested with *Bgl*II. The resulting 1.6- and 8.5-kb fragments, carrying the yeast and *Tetrahymena* rDNA telomeres respectively, were ligated into pBR322 opened by *Bam*HI digestion, to produce pYt103 (*b*) and pYt14 (*c*). In pYt103, a deletion of pBR322 sequence adjacent to the terminal telomeric sequence had occurred, and the *Bam*HI linker, shown in brackets, was lost. In the procedure used to produce pYt14 (*b*), *Bal*31 nuclease (BRL) was used instead of *S*₁ nuclease, and the resulting blunt end of the 1.6-kb *Bgl*II end fragment of pSZ219 was ligated directly to the *Pvu*II site of the pBR322 vector. Cloning procedures followed ref. 30. Restriction sites: H3, *Hae*III; B2, *Bgl*II; PI, *Pvu*I; B, *Bam*HI; D, *Dde*I; RI, *Eco*RI; P2, *Pvu*II. Parentheses indicate hybrid sites.

telomeric sequence, pYt103 and pYt14, are shown in Fig. 1b. The recombinants were mapped by digestion with restriction enzymes and compared with the restriction maps obtained using pSZ219 purified from yeast cells. The maps were colinear, but the region corresponding to the most terminal ~400 base pairs (bp) of the cloned yeast telomere on pSZ219 was absent from clones pYt103 and pYt14; these data are summarized in Fig. 2a.

The inserts of pYt103 and pYt14 were both sequenced by the Maxam-Gilbert method¹⁹, using the sequencing strategy shown in Fig. 2a. The sequence of the yeast telomere is shown in Fig. 2b. The most noteworthy feature of the sequence is that the last 81 nucleotides, corresponding to the most distal region (with respect to the linear vector) of the subcloned segment in pYt103 consists entirely of tandem repeats of the sequence C₁₋₃A. Like *Tetrahymena* rDNA molecules¹⁷, the yeast telomere on pSZ219 has specific single-strand breaks in the distal ~100 bp of the molecule, allowing nick-translation of this region by *E. coli* DNA polymerase I (ref. 18 and E.H.B., unpublished results). Because this sequence is uninterrupted by G or T residues, it accounts for (1) the observed extensive *in vitro* nick-translation of the yeast telomere in the presence of only the two triphosphates, dATP and dCTP¹⁸, and (2) the specific labelling of the pyrimidine tracts C₃(A), C₂(A) and C(A) which occurs in this *in vitro* reaction (E.H.B., unpublished results). These findings, taken together, strongly suggest that C₁₋₃A repeats extend out to the single-strand breaks. The seemingly irregular repeat unit is reminiscent of the irregular sequence, consisting of tandem repeats of C₁₋₈T, at the termini of *Dictyostelium* rDNA molecules¹¹. However, the yeast telomeric C₁₋₃A repeats, in this and other examples analysed, all conform to a larger repeat unit which can be written: 5'[C₂₋₃A(CA)₁₋₃]'3'. Thus, there appear to be strong constraints, imposed during either their formation or maintenance, on the arrangement of the basic C₁₋₃A repeats.

Chan and Tye²⁰ have described a class of 6.7-kilobase (kb) repeats adjacent to several chromosomal telomeres in yeast. Internal to the C₁₋₃A repeats, the restriction map of the yeast

‡ To whom reprint requests should be addressed.

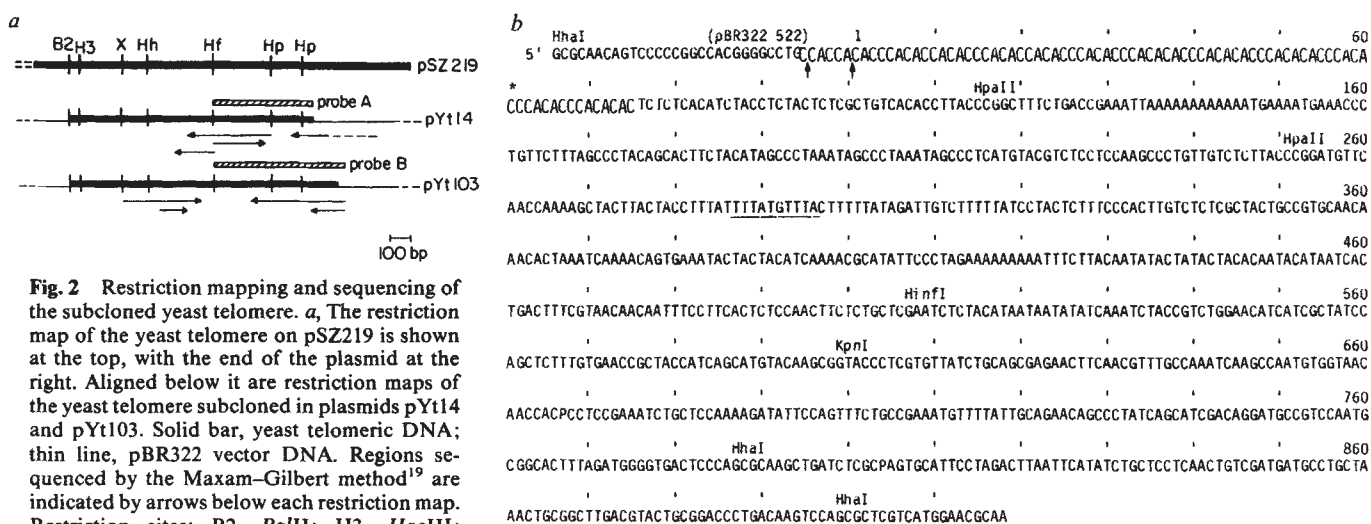


Fig. 2 Restriction mapping and sequencing of the subcloned yeast telomere. *a*, The restriction map of the yeast telomere on pSZ219 is shown at the top, with the end of the plasmid at the right. Aligned below it are restriction maps of the yeast telomere subcloned in plasmids pYt14 and pYt103. Solid bar, yeast telomeric DNA; thin line, pBR322 vector DNA. Regions sequenced by the Maxam-Gilbert method¹⁹ are indicated by arrows below each restriction map. Restriction sites: B2, *Bgl*II; H3, *Hae*III; X, *Xho*I; Hh, *Hha*I; Hf, *Hinf*I; Hp, *Hpa*II. Hatched bars show restriction fragments A and from the distal *Hinf*I site in the telomere to the pYt103. Numbers above the sequence refer to the subcloned telomere, and the last nucleotide in parentheses applies to the pBR322 vector sequence. The asterisk (*) indicates a sequence inferred to have occurred in pYt103 within the subcloned telomere at position 61. The underlined sequence is

telomere cloned in pSZ219 matches a portion of the 6.7-kb repeat. The previously cloned examples of internally located 6.7-kb repeats contained an autonomously replicating segment (ARS)²⁰. The terminally located example we have cloned in pSZ219 has an ARS consensus sequence²¹, TTTATGTTTA (underlined in Fig. 2b), occurring ~700 bp from the molecular end of pSZ219.

To map the telomeres of yeast chromosomes, chromosomal DNA was digested with several restriction enzymes, the digests were fractionated by gel electrophoresis, blotted onto a nitrocellulose filter, then probed with the distal portion of the subcloned yeast telomere. The probe used was ^{32}P -labelled fragment A of pYt14 (Fig. 2a), which lacks a long tract of terminal C_{1-3}A repeats. Figure 3a shows that after digestion with *Hind*III,

*Bam*HI or *Pvu*I, the yeast telomeric sequence in fragment A hybridized strongly with up to six genomic bands, consistent with previous findings¹⁸. In contrast, after digestion with enzymes which cut nearer the yeast terminus of pSZ219 (*Bgl*II, *Xho*I, *Hha*I or *Hinf*I), the predominant hybridization was with a single, heterodisperse genomic band which, as shown below, represents chromosomal termini. The width of this band corresponds to a size range of ± 50 –100 bp. The differences in size between the terminal genomic bands produced using *Bgl*II, *Xho*I, *Hha*I or *Hinf*I were the same as in pSZ219, showing that the internal restriction map of the telomeric segment cloned in pSZ219 was colinear with that of many chromosomal telomeres. However, for each enzyme digest, the mean size of the terminal chromosomal fragment was ~ 100 bp shorter than the

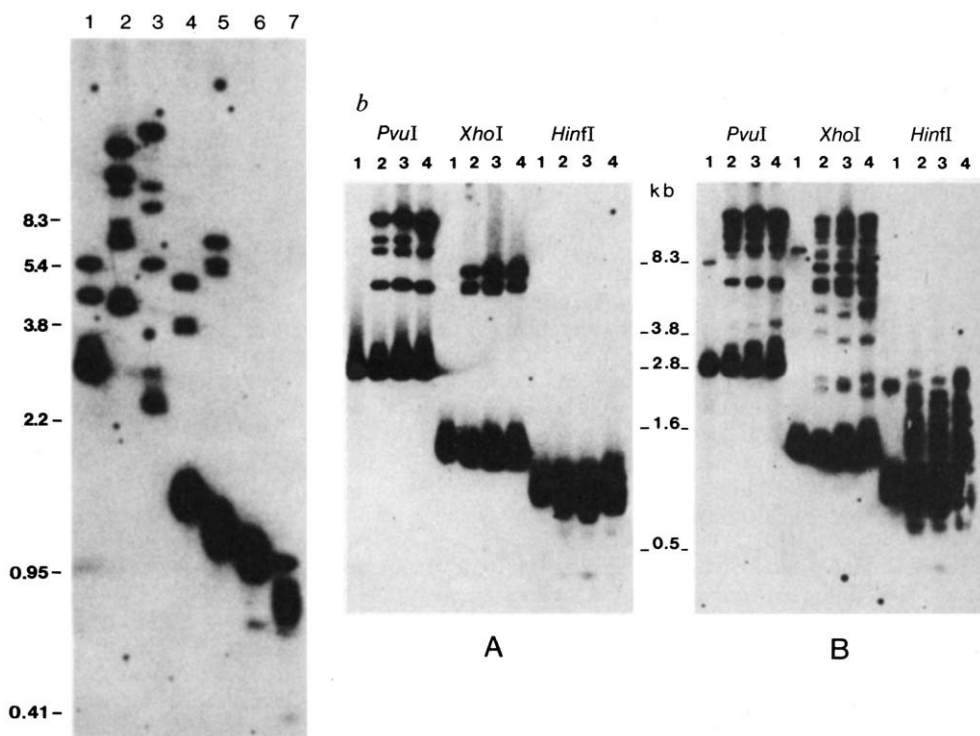


Fig. 3 Hybridization of cloned telomere sequences with genomic yeast DNA. *a*, After digestion with different restriction enzymes and electrophoresis in an agarose gel, a281 DNA was blotted onto a nitrocellulose filter and hybridized with nick-translated ^{32}P -labelled fragment A probe (see Fig. 2a) at 65 °C in 0.6 M NaCl, 0.06 M sodium citrate, pH 7.0 (ref. 12). Restriction enzymes used were *Hind*III (lane 1), *Bam*HI (lane 2), *Pvu*I (lane 3), *Bgl*II (lane 4), *Xho*I (lane 5), *Hha*I (lane 6) and *Hinf*I (lane 7). *b*, Purified pSZ219 plasmid DNA (lanes 1) and genomic DNAs from strains LL20 (lanes 2), a281 (lanes 3) and D2180 (lanes 4) were digested with *Pvu*I, *Xho*I or *Hinf*I as indicated. After gel electrophoresis, duplicate blots were hybridized with ^{32}P -labelled fragment A or fragment B probe (see Fig. 2a), as indicated below each panel.

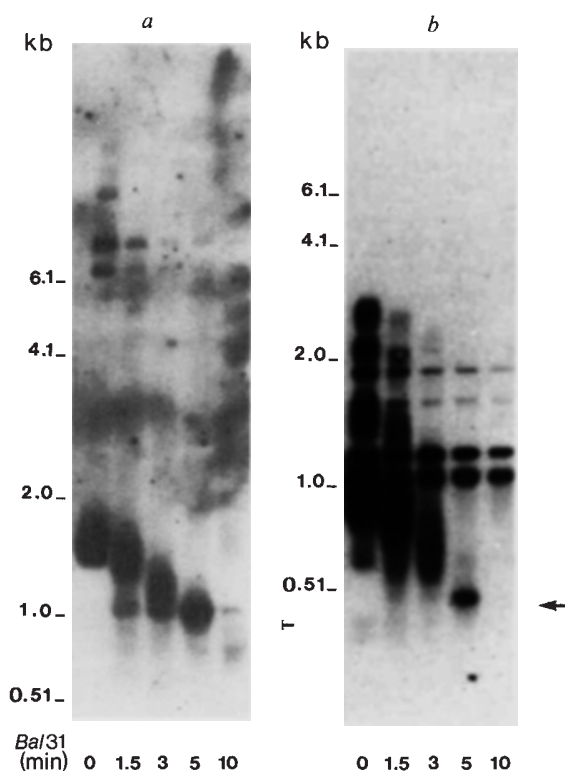


Fig. 4 *Bal31* sensitivity of genomic regions hybridizing with telomeric probes. Genomic DNA from strain D2180 was digested with *Bal31* nuclease for 0, 1.5, 3, 5 and 10 min at 30 °C, in conditions recommended by the manufacturer (BRL), at 5 units of enzyme per 10 µg DNA. The same preparation of *Bal31*-treated DNA was digested with either *XhoI* (a) or *HinfI* (b) and fractionated by electrophoresis in agarose gels, blotted to nitrocellulose filters and hybridized with ³²P-labelled probe A (a) or probe B (b). A pause site, manifested by the sharp band (indicated by arrows) appearing below the terminal heterodisperse fragment as *Bal31* digestion progressed, mapped between positions 119 and 154 in the sequence shown in Fig. 2b.

corresponding terminal pSZ219 fragment, although the chromosomal telomeric band was in each case no more heterodisperse than the band from pSZ219 (Fig. 3b).

To analyse the relative arrangements in the yeast genome of C₁₋₃A repeats and the adjacent telomeric sequence in fragment A, fragment B probe from pYt103 (Fig. 2a) was hybridized with yeast chromosomal DNAs. Probe B is the same as fragment A except that it also contains a long tract of C₁₋₃A repeats. Genomic DNAs from untransformed strain LL20, haploid strain a281 (a derivative of LL20 and the strain in which pSZ219 had been propagated) and diploid strain D2180 were digested with *PvuI*, *XhoI* or *HinfI*, and duplicate blots of each digest were hybridized with either fragment A or fragment B (Fig. 3b). With both probes, only minor differences were observed between the three strains. The fragment A probe hybridized with a subset of the bands hybridizing with fragment B. While over 70% of the hybridization with fragment A was by the common single, prominent, heterodisperse band whose size corresponded to the terminal fragment of the cloned yeast telomere, fragment B probe also hybridized with three additional sharp bands not detected by the fragment A probe (Figs 3b, 4b).

To determine which of the hybridizing genomic fragments were located at the molecular ends of chromosomes, yeast DNA was digested for increasing times with the double-stranded DNA exonuclease *Bal31*, which shortens chromosomal DNAs progressively from their ends. The *Bal31*-treated DNA was then digested with either *PvuI*, *XhoI* or *HinfI*, fractionated by agarose gel electrophoresis, blotted and hybridized with frag-

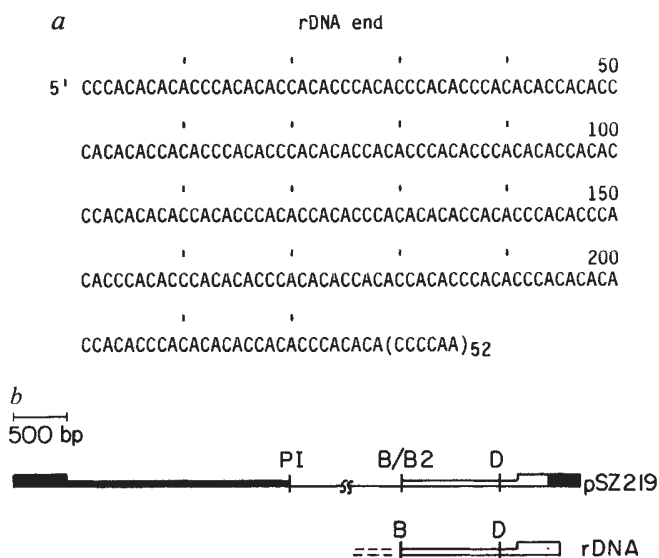


Fig. 5 a, Sequence of the *Tetrahymena* rDNA-derived end of pSZ219-5. Position 1 corresponds to the most distal portion subcloned in prE-Y3 and adjoins the *Bam*HI linker; the direction of increasing numbers is towards the interior of pSZ219-5. b, Summary of the structure of telomeric sequences in pSZ219. The linear plasmid is drawn as in Fig. 1a. Solid bars represent yeast telomeric sequence; very thick portions show the extent of the C₁₋₃A repeat sequence. Open bars represent *Tetrahymena* rDNA sequence; the very thick open bar shows the extent of the CCCCAG repeat sequence in the rDNA purified from *Tetrahymena* cells (one terminal portion shown in lower map), and in pSZ219 (upper map). The extent of yeast C₁₋₃A repeats on the rDNA-derived end is shown by the very thick solid bar.

ment A or B. Figure 4 shows representative results. All initially heterodisperse bands were shortened by *Bal31* digestion; this property identifies them as chromosomal termini. The hybridization signal with the major *Bal31*-sensitive heterodisperse band detected by both probes persisted for over 5 min of *Bal31* digestion, while the hybridization with the three additional heterodisperse bands detected by fragment B but not fragment A was lost after 1.5 min of *Bal31* digestion (Fig. 4b). Thus, we conclude that yeast chromosomal termini, regardless of whether they carry the adjacent fragment A sequence, terminate in C₁₋₃A repeats. We note, however, that C₁₋₃A repeats are not confined to the molecular ends of chromosomes, because fragment B also hybridized with genomic fragments, migrating as sharp bands in gel electrophoresis which were *Bal31*-insensitive (Fig. 4b). Consistent with these results, Walmsley *et al.* have identified C₁₋₃A repeats in internal positions in chromosomes (see accompanying paper²²). These internal C₁₋₃A repeats are not flanked by fragment A sequences.

Preliminary observations on the termini of *Tetrahymena* rDNA molecules in linear yeast plasmids suggested that they had undergone a structural change after their maintenance in yeast. The terminal fragments became longer¹⁸, and they acquired the ability to hybridize (at reduced stringency) with a synthetic homopolymeric probe, poly(dGT), a property shared with the yeast telomeric fragment cloned on the linear vector²³. We subcloned the end of pSZ219 derived from *Tetrahymena* rDNA (Fig. 1). The sequence of the distal subcloned region (between the *DdeI* and *Bam*HI sites shown in Fig. 1c) was determined and is presented in part in Fig. 5a. Although the *Tetrahymena* rDNA sequence was colinear and matched that determined previously for this molecule in *Tetrahymena*²⁴, an additional 228 bp of sequence, consisting entirely of the same type of C₁₋₃A repeat sequence as found at the yeast telomere, was joined directly to the CCCCAG repeats previously shown to occur at the end of the *Tetrahymena* rDNA molecules. The transition from *Tetrahymena*-derived to yeast-derived sequence was abrupt, with no mixing of the two repeat types. The mean

size of the heterodisperse rDNA-derived end fragment of pSZ219 was greater by 200 bp than the mean size of the original heterodisperse rDNA end fragment isolated from *Tetrahymena* cells^{9,17,18}. Therefore, while the *Tetrahymena* sequence was maintained faithfully in yeast, the extreme distal tip of the *Tetrahymena* rDNA molecule was extended, and possibly partly replaced, by at least 228 bp of yeast telomeric C₁₋₃A repeats. The arrangements of telomeric sequences in pSZ219 are summarized in Fig. 5b.

How and why does the addition of telomeric DNA occur? We have shown that *Tetrahymena* telomeres are modified by the addition of yeast telomeric sequences during their replication in yeast. The length heterogeneity of yeast and other telomeres^{9-11,18} suggests that these ends are also being continually modified by the addition and loss of new telomeric repeats. It seems likely that the recently described synchronous lengthening of *Trypanosome* telomeres²⁵ occurs by a similar terminal DNA addition reaction. A DNA addition reaction could account for the genesis of the ends of the macronuclear DNA fragments of ciliates (reviewed in ref. 26). It seems that the telomeric DNA addition reaction may be a universal feature of telomere replication. What then might be the mechanism and function of this reaction?

Any mechanism involving recombination seems unlikely in view of the restricted homology between the *Tetrahymena* and yeast telomeric repeats. In the best match (CCCACA versus CCCAAC), 4 bp of homology are followed by two mismatches. Furthermore, the addition reaction is known to be independent of the yeast repair/recombination gene, *RAD52* (B. Dunn, P. Szauter, M.-L. Pardue and J.W.S., in preparation). The gap-filling model proposed for the growing *Trypanosoma* ends²⁵ is hard to reconcile with the non-palindromic terminal repeats of yeast and all other sequenced telomeres. Finally, the strand-invasion/primer-extension model^{27,28} also seems unlikely in view of the lack of homology between yeast and *Tetrahymena* ends.

We propose that terminal transferase-like activities are responsible for extending the 3' end of the G + T strand of yeast telomeres. Such activities could add single nucleotides, with the final sequence being determined by the substrate specificities of the enzymes involved. Alternatively, preassembled units (such as G₁₋₃T or the longer unit (TG)₁₋₃ TG₂₋₃) could be added. At some point, the growing single-stranded tail of the G + T strand would become a template for DNA primase and DNA polymerase, which could fill-in the C + A strand. The single-strand breaks on the C + A strand could result from incomplete synthesis or the excision of RNA primers. In this model, telomeres are constantly being shortened by incomplete replication and exonucleolytic digestion, but these reactions are in equilibrium with the DNA addition reaction. In this regard, it is interesting that the yeast telomere cloned on the linear plasmid pSZ219 has a slightly longer terminal region than the chromosomally located telomeres, suggesting that the mean length of the telomere is regulated differently in the two cases. The reason that *Tetrahymena* ends, but not random sequences, can function as telomeres in yeast may therefore be that the *Tetrahymena* ends are recognized as a substrate by the yeast DNA addition enzymes. The generation of new telomeres, for example, during macronuclear development in the ciliates²⁶ or the 'healing' of broken chromosomes^{1,3}, may also involve the modification of DNA ends by the telomeric DNA addition enzymes.

We thank J. Thorner for a gift of DNA from strain D2180. This work was supported by grants from the NIH to E.H.B. and J.W.S. J.S. was supported by a University of California Graduate Opportunity Fellowship.

Received 23 January; accepted 10 May 1984.

1. McClintock, B. *Genetics* **23**, 315-376 (1938).
2. Muller, H. J. *Collecting Net* **13**, 181-198 (1938).
3. McClintock, B. *Genetics* **26**, 234-282 (1941).
4. Orr-Weaver, T. L., Szostak, J. W. & Rothstein, R. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6354-6358 (1981).
5. Watson, J. D. *Nature new Biol.* **239**, 197-201 (1972).
6. Cavalier-Smith, T. *Nature* **250**, 467-470 (1974).

7. Bateman, A. J. *Nature* **253**, 379 (1975).
8. Dancis, B. M. & Holmquist, G. P. *J. theor. Biol.* **78**, 211-224 (1979).
9. Blackburn, E. H. & Gall, J. G. *J. molec. Biol.* **120**, 33-53 (1978).
10. Johnson, E. M. *Cell* **22**, 875-886 (1980).
11. Emery, H. S. & Weiner, A. M. *Cell* **26**, 411-419 (1981).
12. Katzen, A. L., Cann, G. M. & Blackburn, E. H. *Cell* **24**, 313-320 (1981).
13. Oka, Y., Shiota, S., Nakai, S., Nishida, Y. & Okubo, S. *Gene* **10**, 301-306 (1980).
14. Herrick, G. & Wesley, R. D. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2626-2630 (1978).
15. Klobutcher, L. A., Swanton, M. A., Donini, P. & Prescott, D. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3015-3019 (1981).
16. Pluta, A. F., Kaine, B. P. & Spear, B. B. *Nucleic Acids Res.* **10**, 8145-8154 (1982).
17. Blackburn, E. H. *et al. Cold Spring Harb. Symp. quant. Biol.* **47**, 1195-1207 (1983).
18. Szostak, J. W. & Blackburn, E. H. *Cell* **29**, 245-255 (1982).
19. Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499-559 (1980).
20. Chan, C. S. M. & Tye, B.-K. *Cell* **33**, 563-573 (1983).
21. Stinchcomb, D. T. *et al. in The Initiation of DNA Replication* (ed. Ray, D. S.) 473-488 (Academic, New York, 1981).
22. Walmsley, R. W., Chan, C. S. M., Tye, B.-K. & Petes, T. D. *Nature* **310**, 157-160 (1984).
23. Walmsley, R. W., Szostak, J. W. & Petes, T. D. *Nature* **302**, 84-86 (1983).
24. Kiss, G. B., Amin, A. A. & Pearlman, R. E. *Molec. cell. Biol.* **1**, 535-543 (1981).
25. Bernards, A., Michels, P. A. M., Lincke, C. R. & Borst, P. *Nature* **303**, 592-597 (1983).
26. Blackburn, E. H. *et al. UCLA Symp. Molec. Cell. Biol.* **20** (in the press).
27. Heumann, J. M. *Nucleic Acids Res.* **3**, 3167-3171 (1976).
28. Goldbach, R. W., Ballen-de Boer, J. E., van Bruggen, E. F. J. & Borst, P. B. B. *Biochim. biophys. Acta* **562**, 400-417 (1979).
29. Din, N. & Engberg, J. *J. molec. Biol.* **134**, 555-574 (1979).
30. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning, a Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
31. Sutcliffe, J. G. *Cold Spring Harb. Symp. quant. Biol.* **43**, 77-90 (1979).

Unusual DNA sequences associated with the ends of yeast chromosomes

Richard W. Walmsley*, Clarence S. M. Chan†, Bik-Kwoon Tye† & Thomas D. Petes*

* Department of Microbiology, University of Chicago, 920 East 58th Street, Chicago, Illinois 60637, USA

† Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, New York 14853, USA

The genome of the yeast *Saccharomyces cerevisiae*, like those of other eukaryotes, contains multiple sequences that hybridize with a poly(GT) probe. We have shown previously that some of the sequences that hybridize with the poly(GT) probe are located near the tips of the yeast chromosomes¹. We report here that many of the remaining poly(GT)-hybridizing sequences are associated with a family of putative replication origins localized near the chromosome ends. These sequences have the general form poly(C₁₋₃A), similar to sequences reported to occur at the tips of chromosomes in the accompanying paper². In addition to poly(C₁₋₃A) tracts, yeast cells contain tracts of alternating C and A bases, similar to those seen in mammalian genomes^{3,4}. These results are used as the basis for a new model of telomere replication.

Approximately one-third of the tracts that hybridize with poly(GT) in the yeast genome are localized at the tips of the yeast chromosomes¹. In order to analyse tracts derived from internal chromosomal sites, we isolated clones that hybridized with poly(GT) from a recombinant plasmid bank containing yeast DNA from the strain DBY 939 (ref. 5). We refer to the poly(GT)-hybridizing sequences as 'poly(CA)' tracts, although data presented in this and an accompanying letter² show that many of these tracts do not contain strictly alternating C and A residues. Of 3,500 clones screened by colony hybridization⁶ with poly(GT), 69 positive clones were detected. If the tracts are not clustered, this result suggests that there is about one tract per 500 kilobases (kb) or 25-30 internal tracts per genome.

To demonstrate whether these poly(CA) tracts were embedded within a previously described repeated yeast sequence, we hybridized poly(GT) with recombinant plasmids containing either ribosomal DNA, the Ty element, the 2 µm yeast plasmid or the X and Y' repeated sequences^{7,8}. We found that the probe hybridized with all plasmids that contained both X and Y', but did not hybridize with other types of repeated yeast genes (data not shown). As for plasmids containing only Y' sequences, poly(GT) hybridized strongly with some and only weakly with others, indicating sequence heterogeneity within the Y' family of repeats.

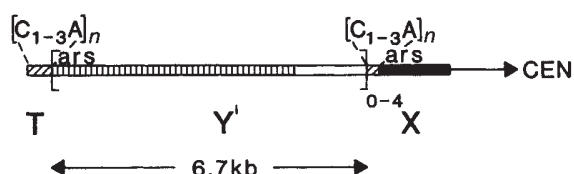


Fig. 1 Organization of X and Y' sequences at the yeast telomere. The organization shown is basically that developed by Chan and Tye⁷. The only change in the diagram from previous depictions is the indication that some yeast telomeres lack Y' sequences. This conclusion is based on Southern blot analysis of yeast telomeres (R.W.W. and T.D.P.; S. Passmore and B.-K.T., unpublished observations) and hybridization of Y' to intact yeast chromosomes (C.S.M.C., B.-K.T. and D. Schwartz, unpublished observations). Although some telomeres may have as many as four Y' sequences, most chromosomal ends have a single Y'. ■, X; ▨, poly(C₁₋₃A); □, the portion of Y' originally designated 131; and ▩, the portion of Y' originally designated Y. As some Y' sequences hybridize with poly(GT), in some cases the poly(C₁₋₃A) tract may extend into Y'.

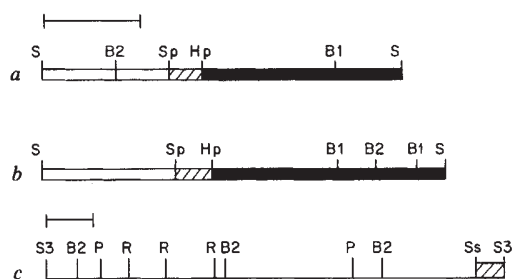


Fig. 2 The location of sequences hybridizing with poly(GT) in plasmids used for sequence analysis. The bars in the diagrams represent 1 kb and the cross-hatched portion of the figure indicates the poly(CA) tract. *a*, pCD2. This plasmid was constructed by inserting yeast DNA that had been partially digested with *Sau*3A into the *Bam*HI site of YEp24 (ref. 5). The 3.7-kb *Sal*I fragment was the only portion of the insertion mapped in detail. *b*, YRp120. The construction of this plasmid has been described previously^{7,8}. The *ars* activity of the plasmid maps between the *Hpa*I and *Bgl*II sites^{7,8}. *c*, pCD9. This plasmid was from the same clone bank as pCD2. The restriction enzymes used in the mapping were: *B*l, *Bgl*II; *B*2, *Bgl*II; *H*p, *Hpa*I; *P*, *Pst*I; *R*, *Eco*RI; *S*, *Sal*I; *S*3, *Sau*3A; *S*s, *Sst*I; and *Sp*, *Sph*I. Other symbols as in Fig. 1. The plasmid pCD9 shows no detectable similarity (independent of the poly(CA) tracts) to X or Y'.

The repeated gene families X and Y' (formerly called 131-Y) were originally characterized as repeated sequences that allowed autonomous replication of plasmids in yeast and were localized near the end of the yeast chromosome^{7,8}. These sequences are often associated with one another in an arrangement shown schematically in Fig. 1. The X repeats are more heterogeneous in both size (0.3–3.75 kb) and sequence relative to the Y' repeats (6.7 kb)^{7,8}. X and Y' show little or no sequence similarity. Although the exact number of X and Y' sequences is unknown, there are at least 10 members of each class per haploid genome⁷. We believe, therefore, that many of the yeast telomeres have the structure shown in Fig. 1.

The portion of the X-containing clones that hybridized poly(GT) was mapped in 10 different recombinant plasmids: nine plasmids described in a previous study^{7,8} (120, 131C, 131M, 131R, 131A, 206, 131J, 131N and 131) and one isolated for this study (pCD2). In all cases, poly(GT) hybridized with the junction between X and Y'. The *ARS* activity (corresponding to a putative replication origin) is separable from the poly(GT) hybridization in plasmids 120, 131C, 131M, 131R, 131A and 206. The positions of poly(GT) hybridization for two of the plasmids (YRp120 and pCD2) are shown in Fig. 2*a, b*.

Sequences of the two tracts shown in Fig. 2*a, b* are shown in Fig. 3. We found that the poly(CA) tract varied in length in the two plasmids (52 base pairs (bp) in pCD2 and 139 bp in

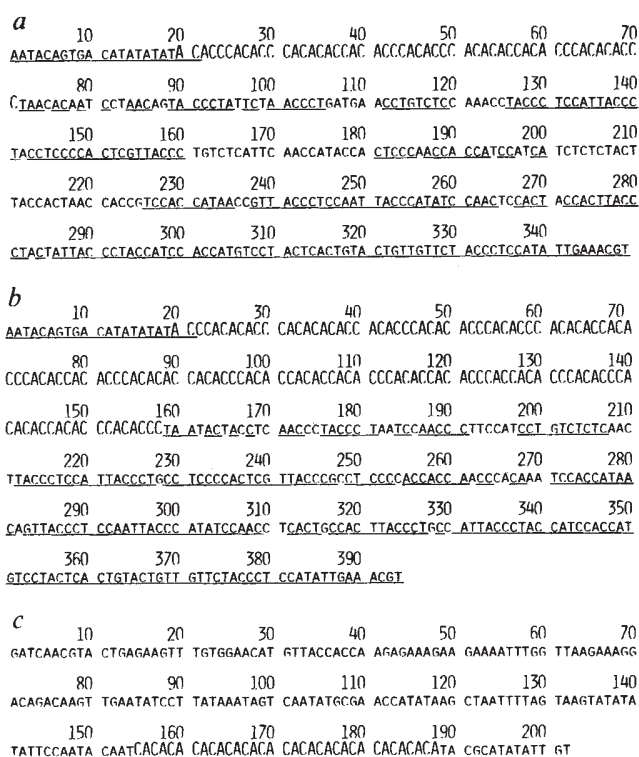
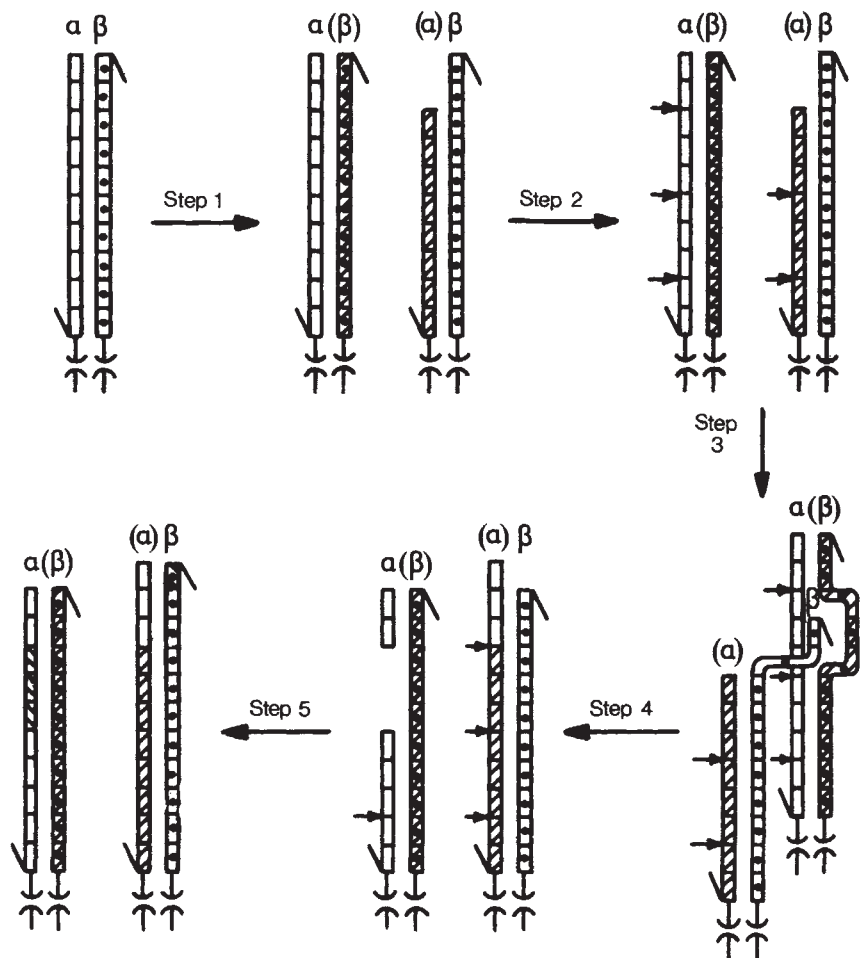


Fig. 3 DNA sequences of regions hybridizing with poly(GT). The recombinant plasmids pCD2(*a*) and YRp120 (*b*) represent X–Y'-containing insertions. For both plasmids, both DNA strands between the *Sph*I and the *Hpa*I sites were sequenced²¹. The 5' strand, starting at the *Sph*I site, is shown. Underlining indicates similarity between the two sequences. The sequence in *c* is from pCD9. A portion of a 750-bp *Sau*3A fragment containing the tract was sequenced in both strands. The 5' strand, starting at the *Sau*3A site, is shown. For all three sequences, the poly(CA) tract is shown by large capitals.

YRp120). For both tracts, however, similar rules apply: no G or T residues interrupt the tract, no pairs of A residues are found and there are never more than three C residues in a row. Thus, the tract can be represented as poly(C₁₋₃A) but is rather more regular than this representation implies: for example, more than 80% of the tract is represented by non-overlapping repeats of the sequence CCACAC. It is presumably the large number of short stretches of alternating CA residues that allows the poly(C₁₋₃A) tracts to hybridize with poly(GT). The sequences that flank the poly(CA) tracts on the Y' side of the boundary are conserved in the two plasmids; the sequences on the other side of the tract have long regions of homology interspersed with regions of non-homology. In particular, the plasmid pCD2 has two insertions (20 and 23 bp) not found in the YRp120 sequence. The poorer conservation of sequence similarity at the X boundary is expected from previous restriction maps of the X sequences^{7,8}. A striking feature of the sequence at the X boundary is the lack of G residues in the strand with the poly(CA) tract (13 G residues in 235 residues for YRp120 and 14 Gs in 278 residues in pCD2).

Szostak and Blackburn, by isolating sequences that allow maintenance of linear plasmids in yeast, obtained cloned yeast telomeres⁹. Shampay *et al.*² have found that the sequence of the tract at the tip of the cloned telomere is essentially the same as those shown in Fig. 3*a, b*. Sequences that hybridize with poly(GT) are added to the ends of the *Tetrahymena* ribosomal DNA when these molecules are cloned in yeast¹. Shampay *et al.*² have shown that these added sequences also have the form poly(C₁₋₃A). It is particularly interesting that the poly(C₁₋₃A) tract of pCD2 (minus the last two base pairs) is identical to a portion of the tract added to the ends of the *Tetrahymena* DNA. The location of this 50-bp identity begins 121 bp from the

Fig. 4 Model of telomere replication. Each telomeric DNA strand is represented by a tandem array of small rectangles. Each rectangle indicates a short repeated DNA sequence. For example, the symbol $\square$ could indicate a (G_xT_y) repeat and the symbol $\square$ could represent a (C_xA_y) repeat. The strands with $\square$ and $\square$ represent DNA synthesized by semi-conservative replication. The repeats $\square$ and $\square$ represent DNA synthesized by repair synthesis. Single-strand nicks are indicated by horizontal arrows and vertical arrows point in the 3' direction on the DNA strand. Step 1: The α and β strands are replicated by a DNA polymerase that has a 5' to 3' elongation activity and requires a primer. Excision of the primer produces a free 3' end in one of the daughter molecules. Step 2: The strand with the C_xA_y repeats is nicked on both DNA molecules, possibly by a site-specific endonuclease. Step 3: The protruding 3' end of the $(\alpha)\beta$ molecule invades the $\alpha(\beta)$ molecule, displacing the (β) strand; this invasion occurs when the molecules are aligned unequally. Step 4: As the α strand is nicked, pairing between the invading β strand and an α fragment followed by dissociation of the two duplexes yields an $\alpha(\beta)$ molecule with a gap in the α strand and an $(\alpha)\beta$ molecule with a 5' protruding strand. Step 5: The gaps are filled by 5' to 3' repair synthesis and nicks are sealed by DNA ligase. As the mechanism shown in this figure results in growth of the yeast telomeres each generation (which is not observed), we suggest that the growth is balanced by exonucleolytic degradation of the exposed 3' terminus.



junction of C_4A_2 *Tetrahymena* sequences with poly($C_{1-3}A$) and extends 50 bp from this junction towards the end of the chromosome.

Although the tracts are conserved between the X-Y' junction and the tip of the chromosome, the sequences that flank the X-Y' tract show no obvious similarity to the sequence adjacent to the tract at the tip². One interpretation of this result is that the only sequence homology between X and the telomeres is the tract itself. We prefer the alternative explanation that the lack of homology represents divergence between different telomeres because DNA sequence analysis of the poly(CA) tracts, together with the information presented in Fig. 1, suggests that yeast has at least two types of telomeres: one type has X sequences and a poly($C_{1-3}A$) tract; the other has one or more Y' sequences integrated within the poly($C_{1-3}A$) tract.

Excluding the poly(CA) tracts located at the tips of the chromosomes, we estimate that the tracts at the X-Y' junctions represent slightly less than half the poly(CA) tracts in the yeast genome. We base this estimate on two experiments. When an X sequence is used as a hybridization probe, about one-quarter (16 of 69) of the clones that hybridize with poly(GT) also hybridize with X. Second, in Southern blots in which 30 non-telomeric bands hybridizing with poly(GT) are detected, only 12 hybridize with an X probe (data not shown). We performed restriction analysis (Fig. 2c) and a sequence analysis (Fig. 3c) of one plasmid (pCD9) that hybridized with poly(GT) but failed to hybridize with X sequences. The poly(CA) tract in this plasmid contains 17 pairs of alternating C and A residues. This strict alternation is characteristic of poly(CA) tracts seen in higher eukaryotes^{3,4}. The sequences flanking the tract have no obvious similarity to those observed in X nor is there a short direct repeat flanking the tract. By Southern hybridization experiments using the DNA flanking the tract as a probe, we found that the poly(CA) tract of the pCD9 is embedded within single-copy DNA (data not shown).

Thus, yeast has both poly($C_{1-3}A$) tracts and co-polymeric CA tracts. The most important question raised by these findings is: what is the functional role of poly(CA) tracts in the cell? One obvious possibility is that the poly($C_{1-3}A$) repeats are required for the maintenance of eukaryotic telomeres.

Watson¹⁰ pointed out that the replication of linear DNA molecule by a DNA polymerase that had 5' to 3' elongation activity and required a primer resulted in one daughter molecule with a 3' protruding end. Failure to complete synthesis of this end would result in a progressive shortening of the chromosome. Models designed to circumvent this problem involve either formation of concatamers¹⁰, palindromes located at the chromosome end¹¹ or formation of a hairpin at the telomere¹². For eukaryotic chromosomes, the hairpin model has received the most attention. Although there is evidence that some viruses have hairpin termini (for example, vaccinia virus¹³), no compelling evidence for such termini exists for eukaryotic chromosomes. In addition, this model for telomere replication fails to explain observed features of eukaryotic telomeres: (1) the telomeres contain short repeated sequences^{14,15}; (2) often one strand of the telomere is extensively nicked^{9,14}; (3) DNA sequences resembling the yeast telomeric sequences are added to the ends of the *Tetrahymena* and *Oxytricha* telomeres when these molecules are cloned in yeast^{1,2,16}; and (4) telomeres derived from a single DNA molecule can be heterogeneous in length⁹. Although some of these observations can be made consistent with the hairpin model by additional assumptions¹⁷, none is explained easily by this model.

One feature that all the models mentioned above have in common is that the DNA sequences involved in completing telomere synthesis are encoded as part of the telomere. As novel DNA sequences are added to the ends of *Tetrahymena* telomeres when these ends are cloned in yeast^{1,2}, we believe that yeast telomeres are more likely to complete replication by addition of sequences that are not part of the original telomere. One

mechanism for the addition of such sequences is provided by an enzyme that is able to add poly(CA) tracts to the 5' end of incomplete telomeres¹. An alternative possibility is shown in Fig. 4; in this model, replication of telomeres is completed by using repeated sequences derived from other telomeres or from the internally located poly(CA) tracts at the X-Y' boundary. We prefer this model to the one involving template-independent enzymatic addition of poly(CA) because it does not involve an unknown enzymatic activity and it more easily explains the similarity between the X-Y' poly(CA) tracts and the tracts added to the ends of *Tetrahymena* rDNA. In addition, our model suggests a role for the X-Y' poly(CA) sequence as a reservoir of sequences to allow completion of telomere replication. Although several other models have been proposed in which

recombination events are invoked to allow replication of the telomere¹⁸⁻²⁰, our model is better able to explain the observed features of the yeast telomeres.

We thank Carol Deshich and Dorothy Stamenkovich for help in isolating the poly(CA) clones, J. Shampay, E. Blackburn and J. Szostak for sequence data prior to publication, and Drs R. Farber, J. Gall, D. Carrol and E. Blackburn for their comments. The poly(CA) tract in pCD9 was sequenced in Ray Gesteland's laboratory during a sabbatical of T.D.P., who thanks the laboratory staff for help. The research was supported by grants from NIH (GM24110 and R.C.D.A. 1 K04 AG00077 to T.D.P., AI14980 and GM26941 to B.-K.T.), NCI (CA19265 to T.D.P. and R.W.W.) and Bristol-Meyers (to C.S.M.C.). R.W.W. is a Charles W. Webster and a Jules J. Reingold Fellow.

Received 23 January; accepted 30 April 1984.

- Walmsley, R. W., Szostak, J. W. & Petes, T. D. *Nature* **302**, 84-86 (1983).
- Shampay, J., Szostak, J. W. & Blackburn, E. H. *Nature* **310**, 143-157 (1984).
- Miesfeld, R., Krystal, M. & Arnheim, N. *Nucleic Acids Res.* **9**, 5931-5947 (1981).
- Hamada, J. & Kakunaga, T. *Nature* **298**, 396-398 (1982).
- Carlson, M. & Botstein, D. *Cell* **28**, 145-154 (1982).
- Grunstein, M. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3961-3965 (1975).
- Chan, C. S. M. & Tye, B.-K. *Cell* **33**, 563-573 (1983).
- Chan, C. S. M. & Tye, B.-K. *J. molec. Biol.* **168**, 505-523 (1983).
- Szostak, J. W. & Blackburn, E. H. *Cell* **29**, 245-255 (1982).
- Watson, J. D. *Nature new Biol.* **239**, 297-301 (1972).
- Cavalier-Smith, T. *Nature* **250**, 467-470 (1974).

- Bateman, A. J. *Nature* **253**, 379-380 (1975).
- Baroudy, B. B., Benktesan, S. & Moss, B. *Cell* **28**, 315-324 (1982).
- Blackburn, E. H. & Gall, J. G. *J. molec. Biol.* **120**, 33-53 (1978).
- Boswell, R. E., Klobutcher, L. A. & Prescott, D. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3255-3259 (1982).
- Pluta, A. F., Dani, G. M., Spear, B. B. & Zakian, V. A. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Bernards, A., Michels, P. A. M., Lincke, C. R. & Borst, P. *Nature* **303**, 592-597 (1983).
- Heumann, J. M. *Nucleic Acids Res.* **3**, 3167-3171 (1976).
- Goldbach, R. W., Ballen-de Boer, J. E., Van Bruggen, E. F. J. & Borst, P. B. B. *Biochim. biophys. Acta* **562**, 400-417 (1979).
- Van Der Ploeg, L. H. T., Liu, A. Y. C. & Borst, P. *Cell* **36**, 459-468 (1984).
- Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499-559 (1980).

Triple-point behaviour of human haemoglobin

Stanley J. Gill & Brough Richey

Department of Chemistry, University of Colorado, Boulder, Colorado 80309, USA

The recent crystallographic observation of Brzozowski *et al.*¹ demonstrates the existence of half-oxygenated haemoglobin crystals. At first sight this seems to imply the discovery of a unique molecule, defined by exactly two bound oxygens. However, the formation of crystals with a specific degree of ligation is predicted to occur when two distinct crystalline phases coexist, namely at the triple point². It seems likely that the crystals studied by Brzozowski *et al.*¹ were selected from a mixture of T and R crystalline forms obtained in conditions where the triple point existed. In general, the equilibria between haemoglobin solid and liquid phases are governed by the chemical potential of a control ligand and the ligand-binding properties of each phase. We report here that a unique situation arises when there are two different crystalline phases, each with characteristic oxygen-binding properties. One can then predict the existence of a triple point, defined by a specific oxygen partial pressure, where both solid phases coexist². At this point the degree of oxygen saturation in each solid phase is uniquely specified. These considerations could explain the specifically ligated crystals found by Brzozowski *et al.*¹.

An early observation by Perutz, cited by Wyman³, notes the likelihood of triple-point behaviour. Such effects have recently been observed in solubility studies of various human haemoglobins⁴. Binding curves for the crystalline phases allow one to define the oxygen pressure at the triple point and to predict the amount of oxygen bound to the solid T form. Direct measurements of oxygen binding to HbS fibres have been obtained by Sunshine *et al.*⁵. Oxygen-binding curves for the solid phases had been determined indirectly from a study of the solubility of haemoglobin at different oxygen pressures⁴. In both of these studies it was found that solid-phase binding is non-cooperative and can be described with a single binding constant.

We are interested in describing the binding of oxygen to crystals of haemoglobin A (HbA) and thus wish to obtain the relation between the degree of oxygen saturation in the crystals and the oxygen partial pressure. Unligated haemoglobin crystals

are characterized by the T state and added oxygen will therefore initially bind to solid haemoglobin in the T form. The binding curve for T-form crystals, determined from the results of Haire *et al.*⁴ on haemoglobin A in polyethylene glycol (PEG), is shown in Fig. 1. When the oxygen partial pressure reaches that of the triple point, a new crystal form (R) appears. The R crystal form has a much higher oxygen affinity⁴ than the T form, as shown by the binding curve in Fig. 1. Thus, addition of oxygen at the triple point causes the conversion of the less oxygenated T-form crystals to the more oxygenated R-form crystals to proceed until all T-form crystals are eliminated. As the oxygen pressure is fixed at the triple point, this conversion process yields the vertical line seen between the R and T binding curves. This process is entirely analogous to the conversion of ice to liquid water at its melting point, where the amount of liquid and solid phases depends on the amount of heat added.

The stability of the R and T solid phases is seen by calculating the chemical potentials of the unligated haemoglobin species (μ_M) in the two solid phases as a function of oxygen partial pressure. These chemical potentials involve a standard-state term for each form (we take the standard-state value for the T form to be zero) and the term $RT \ln \alpha_0$ where α_0 is the fraction

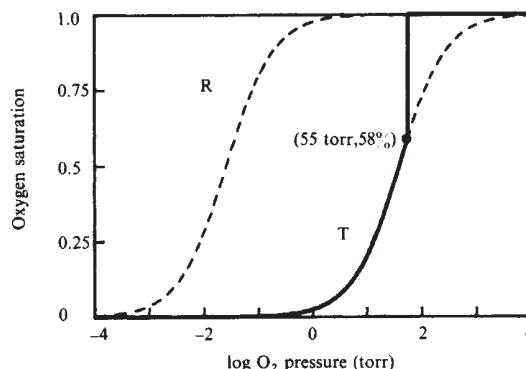


Fig. 1 Binding curves for oxygen to haemoglobin A in R and T crystalline forms. These are single-constant curves with $K_T = 0.026$ torr⁻¹ and $K_R = 39$ torr⁻¹ as determined from solubility measurements in 17.5% PEG at 25 °C, pH 7.0, 0.1 M phosphate buffer⁴. The solid line denotes the actual binding behaviour to the solid phases and the vertical segment represents the triple-point region. Both crystalline forms are present in the triple-point region—note that the T form is 58% saturated at this point.

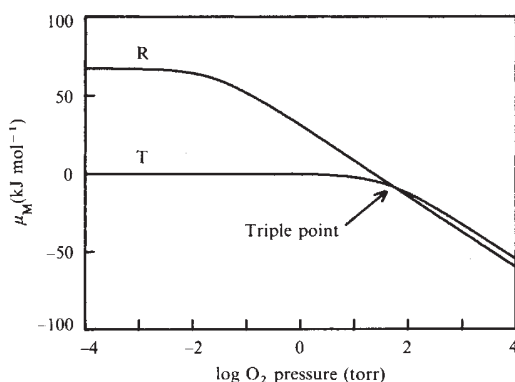


Fig. 2 The chemical potentials of the unligated haemoglobin tetramers in T and R crystalline forms as a function of the logarithm of the oxygen partial pressure. The curves shown are the result of calculations using the parameters of Fig. 1 and assuming that the standard-state chemical potential of the T form is zero.

of unligated haemoglobin in a given solid phase. The difference in the standard-state values of the chemical potentials for the two phases may be evaluated at the triple point where the chemical potentials are equal. Figure 2 shows the chemical potential of HbA as a function of the logarithm of oxygen pressure. The chemical potentials for other species with specific oxygen stoichiometry, such as HbO_2 , $\text{Hb}(\text{O}_2)_2$, etc., may be obtained by adding the stoichiometrically weighted oxygen chemical potential to that of the unligated species. Figure 2 shows that the T form is the most stable form at oxygen pressures below the triple point whereas the R form is the most stable form at pressures above the triple point.

The partial pressure of oxygen at the triple point can be calculated from the results of Haire *et al.*⁴. This calculation uses the condition that both crystalline forms at this point have the same solubility, which is determined by the ratio of the binding polynomials for the liquid and solid forms⁴. The triple point is calculated to be 55 torr for HbA in these conditions; at this pressure the solid T form is 58% saturated and the solid R form is 99.9% saturated. The fact that the saturation value calculated for the T form at the triple point agrees well with the half-saturated T-state crystals found by Brzozowski *et al.*¹, strongly suggests that their crystals were obtained at the triple point. Their crystallographic information allowed them to specify that only the α -chain binding sites were occupied.

We also wish to point out how the binding curve of solid haemoglobin shown in Fig. 1 has a bearing on the likelihood of obtaining crystals with a specific degree of ligation. If a liquid solution of haemoglobin, saturated anywhere between 58 and 99.9%, is completely crystallized, the resulting mixture will contain crystals of both R and T forms. In these conditions the specific degree of oxygenation of the crystals will be 58% for the T form and 99.9% for the R form. Thus, a wide range of oxygen saturation in the liquid phase can give rise to a mixture of crystals, each with a well defined degree of ligation such as those found by Brzozowski *et al.*¹.

It should be stressed that the location of the triple point and thus the corresponding degree of saturation of the R and T crystals will in general depend on specific solution conditions and on the species or mutant form of haemoglobin. The existence of a triple point suggests that there are only two unique crystalline forms for haemoglobin and supports the general features of the allosteric mechanism for oxygen binding.

This work was supported by NIH grant HL22325.

Received 7 February; accepted 3 May 1984.

1. Brzozowski, A. *et al.* *Nature* **307**, 74–76 (1984).
2. Wyman, J. & Gill, S. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5239–5242 (1980).
3. Wyman, J. *Adv. Protein Chem.* **19**, 280 (1964).
4. Haire, R. N. *et al.* *Biochem. biophys. Res. Commun.* **101**, 177–182 (1981).
5. Sunshine, H., Hofrichter, J., Ferrone, F. A. & Eaton, W. A. *J. molec. Biol.* **158**, 251–273 (1982).

Spontaneous generation and amplification of optical activity in α -amino acids by enantioselective occlusion into centrosymmetric crystals of glycine

I. Weissbuch, L. Addadi, Z. Berkovitch-Yellin, E. Gati, M. Lahav & L. Leiserowitz

Department of Structural Chemistry, The Weizmann Institute of Science, Rehovot 76100, Israel

Recent studies of crystal growth in the presence of specific additives have suggested a new approach to the generation and amplification of optical activity in initially racemic mixtures^{1–4}. This approach is based on the observation that particular enantiomers can be selectively incorporated not only into chiral crystals^{5–7} but also into the opposite enantiotopic faces of centrosymmetric crystals^{3,4,8}. If one such face of a centrosymmetric crystal is blocked, therefore, enantioselective occlusion into the opposite face will remove one enantiomer preferentially from the racemic mixture. We report here that this process can take place with centrosymmetric crystals of glycine, in a way that may be of relevance to the generation of optical activity during prebiotic evolution. Glycine crystals float in solution, such that only one face is available for growth. The addition of a particular enantiomer of a hydrophobic amino acid orientates the floating crystals in such a way that the face that occludes amino acids of the opposite chirality is exposed to the solution. Thus, a statistical fluctuation in the orientation of crystals floating in an initially racemic solution could lead to the generation and amplification of optical activity.

When glycine is crystallized from pure aqueous solution, it appears in the monoclinic, centrosymmetric α -form⁹ (space group $P2_1/n$; Fig. 1a) with a bipyramidal morphology⁴ (Fig. 1b). Crystallization in the presence of other, resolved α -amino acids of, say, D-configuration yields pyramidal crystals with an (010) basal face (Fig. 1c); by symmetry the L-amino acids induce an (0 $\bar{1}$ 0) basal face (Fig. 1d). Aqueous solutions of glycine containing racemic mixtures of other α -amino acids yield {010} crystal plates⁴ (Fig. 1e). We further demonstrated that on occlusion during the last process, the α -amino acids segregate enantioselectively at the opposite sides of the plate-like crystal along the unique *b*-axis. During growth the D-amino acids, in amounts of up to 0.3%, are occluded through the (010) face and the L-amino acids are occluded through the enantiotopic (0 $\bar{1}$ 0) face, the two faces being related by a centre of inversion and a glide plane. This resolution has been explained in terms of the orientation of the glycine molecules at these faces; only a D-amino acid molecule can readily substitute glycine at the growing (010) face, and then only certain glycine molecules can do this (shown unshaded in Fig. 1a), because in such a substitution the side chain of the D-amino acid would emerge out of this face (in the glycine C-H₁ direction). The adsorption of an L-amino acid at the same site is precluded because its side chain would have to lie on the crystal surface (in the C-H₂ direction) and this would involve overwhelming steric hindrance. The other glycine molecules at this face (shown shaded in Fig. 1a), which are related to the abovementioned (unshaded) set by a centre of inversion, cannot be replaced by either D- or L-amino acids, for that would entail penetration of the amino acid side chain into the crystal. On symmetry grounds we conclude that only an α -amino acid molecule of L-configuration can be adsorbed at the (0 $\bar{1}$ 0) face⁴.

The above results suggest that if one of the {010} [that is, (010) or (0 $\bar{1}$ 0)] faces of a single glycine crystal is blocked during growth, the crystal will occlude exclusively the appropriate enantiomer through the other, exposed face. If one starts with a racemic mixture of amino acids, such a process should result in their net resolution, one enantiomer being occluded in the

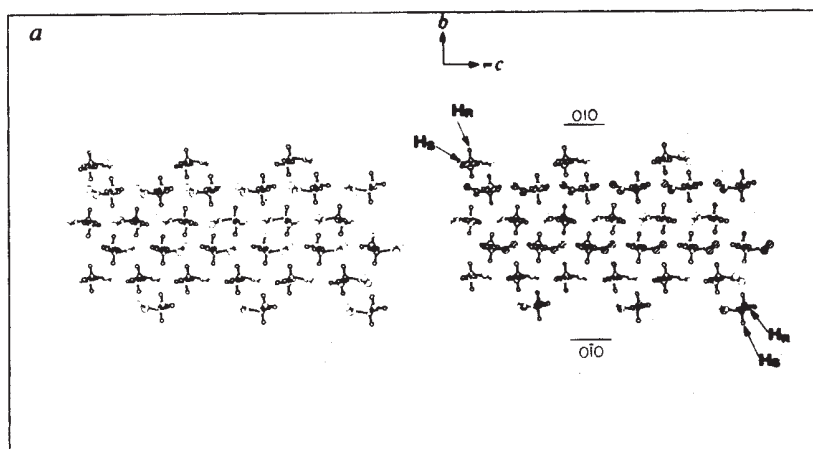
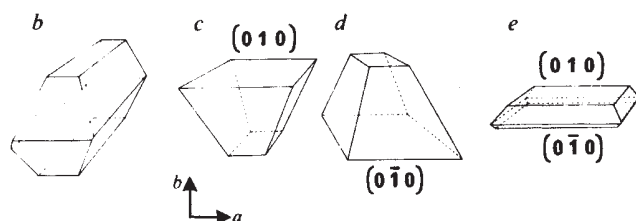


Fig. 1 *a*, The packing arrangement of α -glycine viewed along the *a* axis. The molecules of glycine adopt an asymmetrical configuration in the crystal; the hydrogens H_1 (pro-*S* configuration) of the shaded set of molecules emerge at the $(0\bar{1}0)$ face, whereas the H_1 atoms (pro-*R*) of the unshaded set emerge at the (010) face. *b–e*, Morphologies of α -glycine crystals. *b*, Pure α -glycine crystal grown from water (4.4 M); *c*, glycine grown in the presence of D- α -amino acids; *d*, glycine grown in the presence of L- α -amino acids; *e*, glycine grown in the presence of D,L- α -amino acids.



crystal, with the solution becoming relatively enriched in the other. In the present context, blocking is achieved by growing a single crystal of glycine at an interface, such as water/air or water/glass, taking advantage of the fact that affected glycine crystal plates (density 1.16 g cm^{-3}) have a tendency to float with one of their $\{010\}$ faces exposed to air. Glycine was therefore grown in the presence of racemic mixtures of various amino acids and the resulting plate-like single crystals floating at the surface or lying on the glass bottom were analysed by HPLC for the enantiomeric content of the occluded additive¹⁰. Table

Table 1 Orientation of floating glycine crystals and enantiomeric excess of occluded additives in some typical experiments

No.	Additive (% wt/wt of glycine)	No. of crystals checked	Crystal orientation (face exposed to air)	% Enantiomeric excess of occluded enantiomers
1	D, L-Leu (1–2)	{10 15	{(010) (010)	98 L-Leu 97 D-Leu
2	D, L-Val (1–3)	{15 13	{(010) (010)	95–100 L-Val 100 D-Val
3	D, L-Val (0.5–1) + D, L-Leu (0.5–1)	{5 5	{(010) (010)	100 L-Val + 100 L-Leu 100 D-Val + 100 D-Leu
4	D, L-Val (1) + D, L-Leu (1) + D, L-Glu (1)	4	(010)	100 D-Val + 100 D-Leu + 100 D-Glu
5	D, L-Leu (1) + D, L-His (1)	4	(010)	100 L-Leu + 100 L-His
6	D, L-Phe (0.5)	{5 5	{(010) (010)	100 L-Phe 100 D-Phe
7	D-Leu (1–3)	30	(010) Pyramids	—
8	L-Leu (1–3)	30	(010) Pyramids	—
9	D-Ala (1–5)	30	{(010) Pyramids (010) Plates	—
10	L-Ala (1–5)	30	{(010) Pyramids (010) Plates	—
11	D-Leu (1) + L-Glu (1)	4	(010)	100 L-Glu only
12	L-Leu (1) + D-Glu (1)	4	(010)	100 D-Glu only
13	D-Leu (1–3) + L-Ala (1–3)	10	(010)	—
14	D-Leu (1) + L-Lys (1)	12	(010)	—
15	L- α -aminobutyric acid (1) + D-Ser (1)	5	(010)	—
16	L- α -aminobutyric acid (1) + D-Ala (1)	5	(010)	—
17	L-Phe (1–1.5) + D-Ala (1–3)	10	(010)	—
18	D-Phenylglycine (1) + L-Ala (1)	4	(010)	—
19	D-Phe (1–2) + L-Ala (1–3)	12	(010)	—
20	D-Ala (1–2) + L-His (1–2)	{14 17	{(010) (010)	—
21	D-Ala (3) + L-Ser (3)	{12 15	{(010) (010)	—
22	D-Ser (1–2) + L-His (1–2)	{13 14	{(010) (010)	—
23	D-Leu (1) + D, L-Glu (3)	10	(010)	100 L-Glu only
24	L-Leu (1–3) + D, L-Glu (1–3)	30	(010)	100 D-Glu only
25	D-Leu (1–3) + D, L-Ser (1–3)	30	(010)	—
26	L-Leu (1) + D, L-Tyr (1)	4	(010)	100 D-Tyr
27	L-Phe (0.5) + D, L-Glu (3)	30	(010)	—
28	D-Leu (3) + D, L-His (1.5) + L-Asp (1.5) + D, L-Glu (1.5)	10	(010)	100 L-His, 100 L-Asp, 100 L-Glu
29	L-Leu (1) + D, L-phPG (0.5) + D, L-Met (0.5) + D, L-Glu (1)	12	(010)	100 D-phPG, 100 D-Met, 100 D-Glu
30	D-Leu (1) – D, L-phPG (0.5) + D, L-met (0.5) + D, L-Glu (1)	16	(010)	>95 L-phPG, 100 L-Met, >95 L-Glu

Fig. 2 Plate- and pyramid-shaped crystals of glycine grown in the presence of: *a*, hydrophobic D-amino acids; *b*, hydrophilic D-amino acids, *c*, D-leucine + hydrophilic L-amino acids.

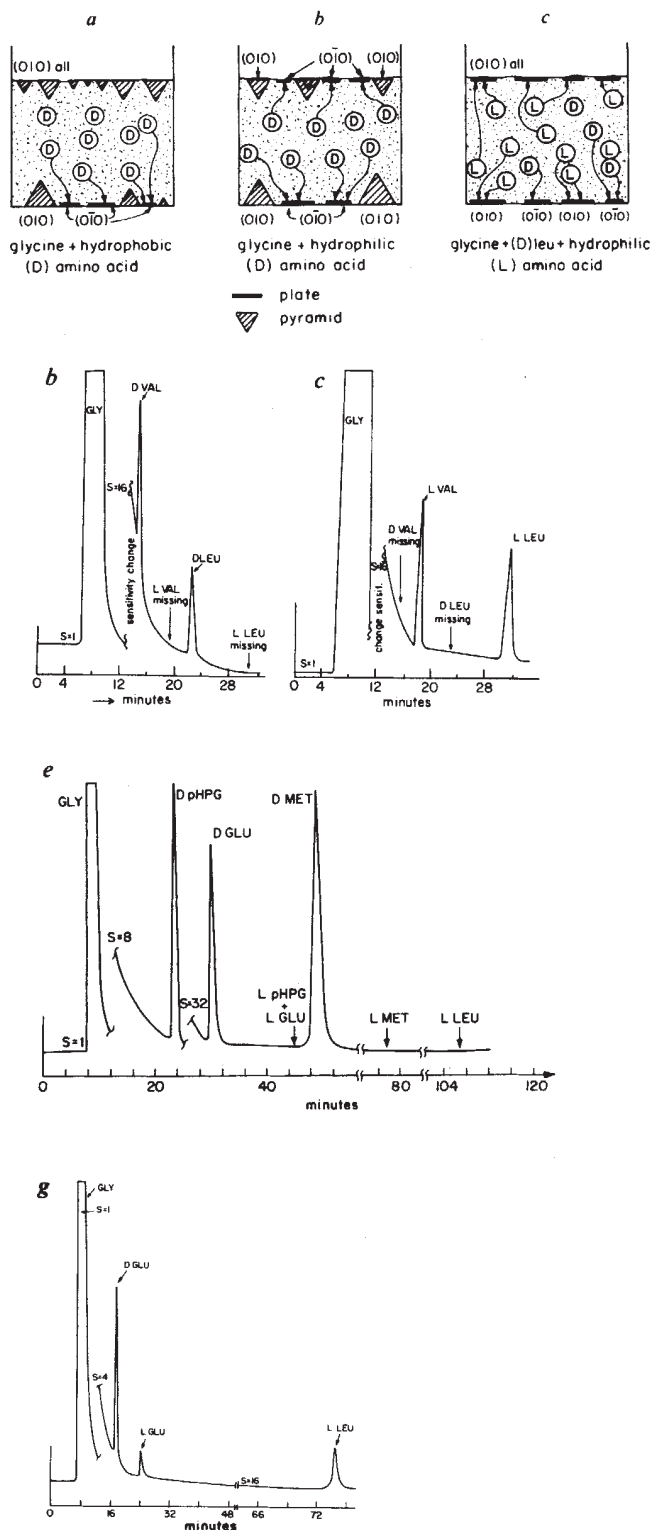
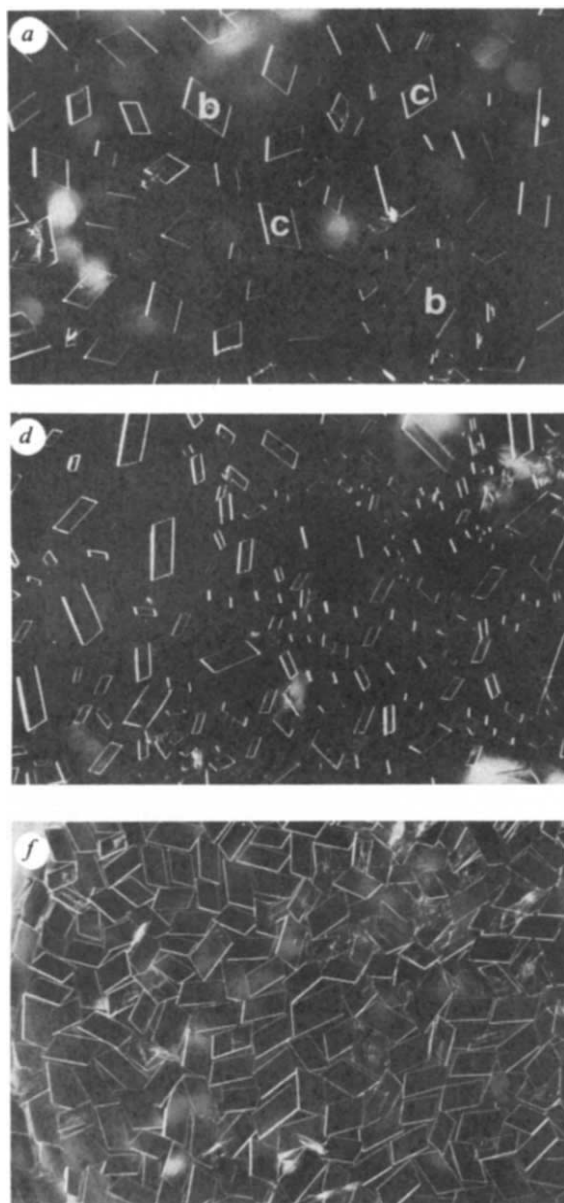


Fig. 3 *a*, Plate-like crystals of glycine floating at an air/water interface, as grown from a water solution (4.4 M) containing racemic 1% wt/wt D, L-Leu and 1% wt/wt D, L-Val. The crystals display both enantiomorphic morphologies. *b*, HPLC analysis of crystals of glycine grown as in *a* with their (010) face exposed to air, after washing with water. Conditions used: reversed-phase column (25 cm $\times$ 4.5 mm) self-packed with 5 μ m Nucleosil C₁₈ (Machery Nagel). Mobile phase composition: aqueous solution of cupric acetate (4 mM) and *N,N*-dipropyl-L-Ala (8 mM) + 7% acetonitrile, at pH 5.3–5.5 (ref. 10); post-column derivatization with *o*-phthalaldehyde and mercaptoethanol and fluorescence detection; flow rate 0.4 cm³ min⁻¹, sensitivities (S) as indicated¹⁰. *c*, HPLC analysis of glycine crystals grown as in *a* with their (010) face exposed to air, after washing with water. Conditions as in *b*. *d*, Plate-like crystals of glycine floating at an air/water interface, grown in the presence of 1% wt/wt L-Leu and racemic *p*-hydroxyphenylglycine (0.5% wt/wt), racemic Glu (1%) and racemic Met (0.5%). All the crystals float with their (010) face exposed to air and exhibit the same enantiomorphic morphology. *e*, HPLC analysis of the enantiomers of *p*-hydroxyphenylglycine (pHPG), Glu, Met and Leu occluded inside the crystals of glycine grown as in *d*. Only the D-enantiomers of pHPG, Glu and Met are found. Conditions: mobile phase without acetonitrile, flow rate 0.4 cm³ min⁻¹. In conditions in which blanks record Leu as low as 1×10^{-4} % of Gly, no L-Leu was detected. *f*, A crystalline crust of glycine grown at an air/water interface. The water solution (4.4 M) contains 1% wt/wt L-Leu and 3% wt/wt D, L-Glu. *g*, HPLC analysis of the crust formed by glycine crystals as shown in *f*, indicating an enantiomeric excess of 70% of occluded D-Glu. Conditions: mobile phase without acetonitrile, flow rate 0.5 cm³ min⁻¹.

1 (lines 1–6) summarizes some typical results. Over 500 experiments showed that, as expected, the D-amino acids were incorporated only into those floating glycine crystals whose (010) faces were exposed to the water solution, while the L-amino acids are incorporated only into the crystals with exposed (0 $\bar{1}$ 0) faces (Fig. 2a–c). HPLC analysis of plates grown on the glass bottoms showed similar resolution, albeit with a lower enantiomeric excess.

Without an amplification mechanism, it is unlikely that a systematic net resolution of the α -amino acid additives would be obtained from crystals grown in this way in a closed system. If, on the other hand, α -amino acids which have been partially resolved in the first random crystallization are able to induce correct orientation in the subsequent flotation of further glycine crystals, asymmetry may be both preserved and amplified. To test this idea, we investigated the ability of resolved α -amino acids to orient glycine crystals at water–air interfaces (see Table 1 legend for the specific conditions for the experiments given below). In these conditions the results were fully reproducible.

We found that resolved hydrophobic α -amino acid additives such as Leu, Phe, Phe-Gly and α -aminobutyric acid of, say, D-configuration, have indeed a pronounced orientating effect; these additives induced the formation of pyramidal crystals all floating with their well developed (010) faces directed upwards (Table 1, lines 7, 8). On the other hand, in the presence of these same additives, crystals of two distinct habits were found on the glass bottom of the vessel—pyramids lying on their (010) face and plates lying on their (0 $\bar{1}$ 0) face (Fig. 2a)—indicating that there is no orientating effect here. Furthermore, in the presence of resolved alanine or other more hydrophilic amino acids, both plate-like and pyramidal floating crystals were obtained in the given range of additive concentrations (Fig. 2b and Table 1, lines 9, 10). Combinations of D-Leu and L-Ala (or L hydrophilic amino acids) in equal concentrations yielded, without exception, parallelepipeds floating with their (010) faces pointing upwards (Fig. 2c and Table 1, lines 11–19). A similar effect was found at a water/hexane interface. Finally, combinations of non-hydrophobic amino acids, (for example, D-Ala and L-Glu or L-Ser) did not lead to preferential orientation of the floating plate-like crystals (Table 1, lines 20–22). The absolute orientations of the floating crystals, that is, (010) or (0 $\bar{1}$ 0) directed upwards⁴, were established by visual observation: their morphology is such that the two enantiomorphic habits are not superimposable in the plane of the interface (Fig. 3a, d, f). The assignments were verified by X-ray diffractometry on several crystals from each batch. According to HPLC analysis of the above crystals, only minute amounts of the orientating hydrophobic additives are occluded through the upward-pointing face. This result indicates that the orientation of the floating crystals occurs at a very early stage of growth.

Based on the above findings, a novel induced resolution of α -amino acids was then performed by crystallization of glycine in the presence of, for example, L-Leu and mixtures of other (D, L) amino acids (Table 1, lines 23–30). The floating crystals (Fig. 3d) were obtained with their (0 $\bar{1}$ 0) faces directed upwards, as deduced by X-ray diffractometry, and, as shown by HPLC analysis, occluded exclusively the D-amino acids through the exposed, growing (010) face; a typical chromatogram is depicted in Fig. 3e. Furthermore, even in conditions where polycrystalline crusts of glycine were formed (Fig. 3f), resolution occurred with an enantiomeric excess higher than 70% (Fig. 3g) in the examples studied.

The above experiments provide a simple model for the generation and amplification of optically active amino acids in prebiotic conditions. Consider, for example, that the first crystallization of glycine from the evolutionary 'soup' gave a crystal floating with its (010) face directed, by chance, upwards. Thus, the (010) face grew in contact with the solution containing racemic mixtures of α -amino acids. During that process the crystal would have occluded the L-amino acids exclusively and thus the water solution would have been enriched in the corresponding D-enantiomers. The hydrophobic D-amino acids pres-

ent in excess in the water solution would, then, have preferentially directed new floating crystals with their (010) faces upwards and these in turn would, during growth, have occluded only L-amino acids from the solution, thus leading to further D-enantiomer enrichment of the solution.

We are now investigating the detailed mechanism of the orientating effect as well as the minimum enrichment required to trigger amplification. Preliminary studies using Leu as an additive have shown that a D/L-Leu mole ratio as low as 55:45, with a total Leu concentration in the range 2.0–2.2% wt/wt of glycine, results in an assembly of perfectly orientated floating, plate-like crystals of glycine.

The principles applied here may be extended to oligopeptides and to other organic and inorganic materials, provided they exhibit the appropriate crystal point symmetry and also develop enantiotopic faces capable of interacting selectively with D- and L-enantiomers.

We thank the Petroleum Research Fund administered by the American Chemical Society and the US/Israel Binational Science Foundation, Jerusalem, for financial support.

Received 25 January; accepted 11 April 1984.

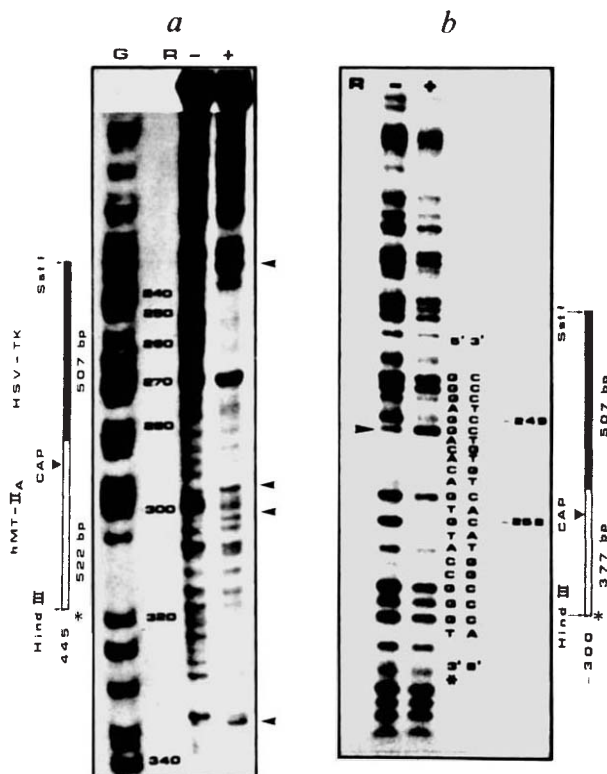
1. Addadi, L. *et al. Nature* **296**, 21–26 (1982).
2. Berkovitch-Yellin, Z., Addadi, L., Idelson, M., Leiserowitz, L. & Lahav, M. *Nature* **296**, 27–34 (1982).
3. Addadi, L. *et al. J. Am. chem. Soc.* **104**, 2075–2077 (1982).
4. Weissbuch, I. *et al. J. Am. chem. Soc.* **105**, 6615–6621 (1983).
5. Green, B. S. & Lahav, M. *J. molec. Evol.* **6**, 99–115 (1975).
6. Addadi, L., Van Mil, J., Gati, E. & Lahav, M. *Origins Life* **11**, 107–118 (1981).
7. Bonner, W. A. & Kavasmanek, P. R. *J. Am. chem. Soc.* **99**, 44–50 (1977).
8. Chenchalah, P. C., Holland, H. L. & Richardson, M. F. *JCS chem. Commun.*, 436–437 (1982).
9. Albrecht, G. & Corey, R. B. *J. Am. chem. Soc.* **61**, 1087–1031 (1939).
10. Weinstein, S. *Angew. Chem.* **21**, 218 (1982).

Erratum

Characterization of DNA sequences through which cadmium and glucocorticoid hormones induce human metallothionein-II_A gene

Michael Karin *et al. Nature* **308**, 513–519 (1984)

IN Fig. 5a, the footprint due to receptor binding was not evident because a poor original was reproduced darkly. The figure is reproduced correctly below. The strong receptor footprint between –265 and –245, as well as the weak footprint around –324, are clearly visible.



Cold Gondwana, warm Tethys and the Tibetan Lhasa block

ALLÈGRE *ET AL.*¹ suggest that there is a climatic and hence palaeolatitude paradox implied by late Carboniferous glacial deposits and early Permian warm Cathaysian floras, both being present in the Lhasa block of southern Tibet. To account for this paradox they claim that, if it formed part of Carboniferous Gondwanaland, the Lhasa block must have detached in the early Permian.

Allègre *et al.*¹ also question, without offering evidence, the interpretation of the Carboniferous Damxung-Linzhu pebbly slates as tillites. I have never seen these deposits but it seems significant that in many of these Asian blocks proposed as parts of eastern Gondwanaland², pebbly mudstone facies of late Carboniferous age have been identified and interpreted as glacial marine diamictites (Fig. 1). Cameron *et al.*³ explain the lithological characteristics of these pebbly mudstone deposits by the reworking and turbiditic redeposition of ice-rafted sub-glacial or fluviglacial debris. They explain the presence of limestone with warm water faunas towards the top of the 'pebbly mudstone' by an amelioration of the climate. However, Smith *et al.*⁴ show that Australian Gondwanaland moved closer to the South Pole from mid-Carboniferous to early Permian times, which would not favour climatic amelioration of eastern Gondwanaland.

The explanation of the climatic paradox may be that during mid-Carboniferous times southern Tibet was well within (<600 km) the cold continent of eastern Gondwanaland. By the postulated removal of central Tibet and Indochina², the Lhasa block became, in early Permian times, part of the newly rifted continental margin of Gondwanaland, suddenly invaded by the warm waters of the new ocean, Tethys II, that was spreading across the early Permian tropics.

The postulated rifting of southern Tibet from Gondwanaland in the mid-Jurassic (Fig. 1) correlates with the evolution of

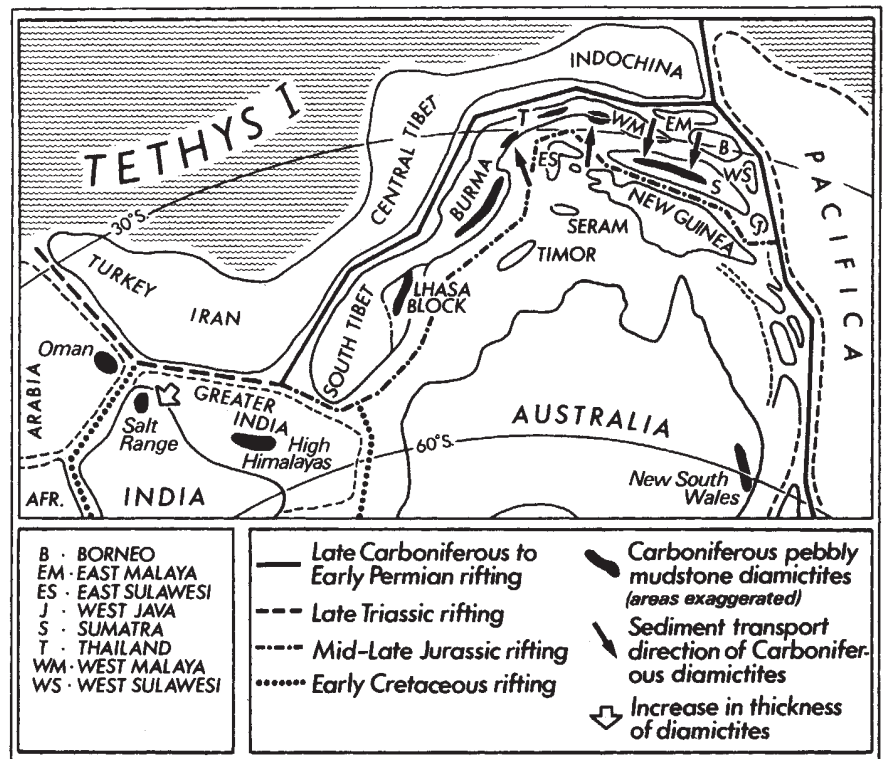


Fig. 1 Correlation of the late Carboniferous marine glacial pebbly mudstones (diamictites) of the eastern part of greater Gondwanaland. The qualitative reconstruction² is based on the Antarctic-India-Australia assembly of ref. 4. Present coastlines or outlines are for reference only. The shapes of the Asian fragments originating in greater Gondwanaland have been modified by tectonic deformation⁷ since they left Gondwanaland. The original shapes have not been reconstructed here but some allowance for crustal shortening has been made. Late Carboniferous marine glacial diamictite data for Sumatra, Malaya, Thailand, Burma, New South Wales, High Himalayas are from ref. 3. Oman glacial marine diamictites are plotted from ref. 8. The western Salt Range diamictites described by Teichert⁹ may be early Permian rather than late Carboniferous. Location of Pacifica and late Carboniferous to Permian and late Triassic rifting are after ref. 2, Jurassic rifting after ref. 5. Early Cretaceous rifting and latitudes are plotted after ref. 4.

the north-west Australia margin⁵. Using the late Cretaceous palaeolatitude of 15°N for the Lhasa block, proposed by Allègre *et al.*¹, implies a spreading rate of ~7 cm yr⁻¹ for 50 Myr if the Lhasa block were rifted from north-west Australia (~15°S) in the mid-Jurassic (170 Myr). That spreading rate compares closely with India following a similar path at 10 cm yr⁻¹ for the first 30 Myr then 5 cm yr⁻¹ for the next 40 Myr (ref. 6).

M. G. AUDLEY-CHARLES

Department of Geology,
University College London,
Gower Street,
London WC1E 6BT, UK

ALLÈGRE *ET AL.* REPLY—Audley-Charles has commented on two distinct aspects of our paper: the Carboniferous versus Permian climatic paradox and the late (late Jurassic) rifting of the Southern

Tibet (Lhasa) block. Before answering these comments, we should point out that the hard core of our observations¹ concerned the central and southern parts of the Lhasa block for periods, in general, after the late Jurassic. Thus Audley-Charles has focused on that part of our data which, both in space and time, certainly needs further study.

Yet, several lines of evidence strongly argue against a late drift of the Lhasa block: (1) The Rhaetian (~200 Myr) flora of the Linzhu basin is identical to the Rhaetian flora from Laurasia and quite distinct from the coeval Gondwanian flora. (2) The ammonoid fauna from the Upper Anisian (~220 Myr) from the Lhasa block seem to be related more closely to those of the European Alps than to those of the Himalayas². (3) The Bangong-Nujiang suture zone which separates southern Tibet (Lhasa block) from northern Tibet (Quantang block) is probably of mid-Jurassic age, as indicated by low-

- Allègre C. J. *et al.* *Nature* **307**, 17–22 (1984).
- Audley-Charles, M. G. *Nature* **306**, 48–50 (1983).
- Cameron, N. R., Clarke, M. C. G., Aldiss, P. T., Aspdon, J. A. & Djinuddin, A. *Proc. Indones. petrol. Ass. 9th A. Conventm* 149–187 (Indonesian Petrology Association, Jakarta, 1980).
- Smith, A. G., Hurley, A. M. & Briden, J. C. *Phanerozoic Paleogeographical World Maps* (Cambridge University Press, 1981).
- Pigram, C. J. & Panggabean, H. *Nature* **301**, 231–234 (1983).
- Molnar, P. & Tapponnier, P. *Scient. Am.* **236**, 30–41 (1977).
- Maddox, J. *Nature* **307**, 9 (1984).
- Williams, G. T. & Whitmore, R. J. *Sci. Technol. Newsl.* **57**, 3 (1983).
- Teichert, C. *Neues Jb. Geol. Paläont., Abh.* **129**, 167–184 (1967).

grade regional metamorphism dated using the U-Pb method on sphenes³. This is consistent with the uppermost Jurassic-lowermost Cretaceous age of the subaerial to shallow marine sediments which unconformably overly the Dongqiao ophiolites⁴. In addition, Jurassic radiolaria have been found in these ophiolites⁵. Finally, the ages of granites associated with this suture range from Upper Jurassic to Lower Cretaceous³. Thus, the Lhasa block must have been accreted to Eurasia before the end of the Jurassic and, if its origin is to be sought amidst austral continents, its separation from Gondwana must have been considerably older. This is also indicated by combining the palaeomagnetic results⁶ and the ages discussed here, along the lines of reasoning in Audley-Charles' last paragraph.

The Palaeozoic history of the Lhasa block is still very poorly known. Indeed, one must explain the simultaneous occurrence of Carboniferous pebbly slates and a Permian Cathaysian flora, in an association which is strongly suggestive of that described in northern Tibet^{7,8}. The cautionary note in our paper¹ was due to the fact that the interpretation of pebbly

slates found in southeastern Asia is controversial, particularly regarding those found in the Phuket Group. They have been interpreted as rift infill⁹, as tillites or, as suggested by Cameron *et al.*¹⁰, as the consequence of reworking and turbiditic redeposition of inrafted subglacial or fluvioglacial debris. On the other hand, one must recall that the Cathaysian flora is traditionally interpreted as being typical of an equatorial climate, indicative of a long period of optimal climate between the Carboniferous and the Permian. Audley-Charles' interpretation implies a progressive change of flora from the South Pole to the northern margin of Gondwana. However, this appears to be contradicted by the discovery of typical Gondwanian flora in the Himalayan Tibet¹¹ which according to Audley-Charles' reconstruction was only about 10°S of the Permo-Carboniferous Lhasa block. Actually, the sole action of warming invoked by Audley-Charles can be replaced or complemented by an early northward migration of the Lhasa block.

The description of Permian brachiopod fauna, which are identical in Australia, western Thailand and New Guinea, argues

strongly for a close geographical connection of the three land masses at the time. On the other hand, available data of Permian fauna from the Lhasa block rather favour the hypothesis of an early northward migration of the block.

CLAUDE J. ALLÈGRE *ET AL.*

*Laboratoire de Géochimie
et Cosmochimie,
Universités de Paris 6 et 7,
4 Place Jussieu, 75230
Paris Cedex 05, France*

1. Allègre, C. J. *et al.* *Nature* **307**, 17–22 (1984).
2. Gu Qingge, He Guoxiang & Wang Yigang *Paleontol. Sinica* 355–356 (1982).
3. Xu, R-H *et al.* *J. Geol.* (submitted).
4. Girardeau, J. *et al.* *Nature* **307**, 27–31 (1984).
5. Wang Naiwen in *Résultats de la coopération franco-chinoise au Tibet (mission 1980)* (eds Mercier, J. L. & Li Guangjin) (CNRS, Paris, in the press).
6. Achache, J., Courtillot, V. & Zhou Yaoxiu, *J. geophys. Res.* (in the press).
7. Bally, A. W. *et al.* *U.S. geol. Surv. Open File Rep.* 80–501 (1980).
8. Norin, E. *Sino-Swedish Expedition Publ.* 29 (Aktiebolaget Thule, Stockholm, 1946).
9. Ridd, M. F. *Nature* **234**, 531–533 (1946).
10. Cameron, N. R. *et al.* *Proc. Indones. petrol. Ass. 9th. a. Conventn.* 149–187 (Indonesian Petrology Association, Jakarta, 1980).
11. Hsü, J. *Scient. Geol. Sinica* **10**, 323–331 (1978).

Oxygen isotope ratios in plants

IN the comprehensive account of their work, Förstel and Hützen¹ state that the different isotope ratios in leaves and fruits on the grape vine are due to their different rates of exchange with the atmosphere. They state that this exchange is more rapid in the leaf than the grape because the leaf has a much greater surface area per unit volume of water. However, gaseous exchange only occurs within the leaf between the cells of the spongy tissues where the anatomical structure is adapted for this purpose, and the grape is not so adapted². The leaf has internal airways with walls that are always wet, allowing gaseous exchange; they do not possess the waxed and cutinized surfaces found on

the exterior of the leaf and fruit. There is no measurable exchange through the outer surface of the mature leaves and fruits—this has been proved by blocking the stomatal pores of the leaves, thus stopping transpiration, etc.

The fruits receive food (sugars dissolved in leaf water) from the leaves via the phloem. In transit, and in the grape itself, this leaf water mixes with the stem water transporting nutrient minerals from the soil through the xylem, thus the final isotope value in the fruit is lower than the average diurnal value of the leaf water (Fig. 2 of ref. 1).

R. PARR

*Miterdale Farm,
Arrington Bridge,
Royston,
Hertfordshire SG8 0AE, UK*

1. Förstel, H. & Hützen, H. *Nature* **304**, 614–616 (1983).
2. Esau, K. *Plant Anatomy*, 422–431 (Wiley, New York, 1953).

FÖRSTEL REPLIES—The idea that gaseous exchange only occurs inside the leaf is not sufficient to describe the fractionation of H₂¹⁸O and H₂¹⁶O between the leaf water and water vapour of the surrounding air. If one takes into account merely internal processes in the leaf, only a simple phase transition from the wet cell walls to a water vapour-saturated gas space must be assumed. The diurnal variation of the relative air humidity surrounding the leaf exterior cannot be expected to have any effect, as has been observed¹. On the other

hand, tracer experiments using enriched water showed that the label from the root zone is transported unaltered through the stem and the stalk²; dilution occurs only in the leaves. Finally, the steady-state tracer concentration in the leaves lies between the ¹⁸O content of the water from the roots and that of the air humidity.

The supply of sugar from the leaves to the fruits is an active process which does not depend on the movement of water. There may be a continuous dilution of the water in the fruit by the inflow of water from the stem or stalk, but the enrichment of H₂¹⁸O (compared with H₂¹⁶O) must be largely due to a water/vapour phase transition between the liquid in the fruit and the surrounding air. Fluid, which moves via the phloem, is diluted rapidly by the root-leaf water flow down to the groundwater level. We have not experimentally observed any influence from an internal leaf-fruit water cycle. Rather, the following conclusion must be valid: the H₂¹⁸O content of grapes remains constant because of the small surface/volume ratio of the fruits, in contrast to the leaves where the H₂¹⁸O content changes quickly during the day. The H₂¹⁸O content of fruit water cannot be due solely to the water supply via the xylem.

H. FÖRSTEL

*Institut für Radioagronomie,
D-5170 Jülich, FRG*

1. Förstel, H. *Radiat. envir. Biophys.* **15**, 323–344 (1978).
2. Förstel, H. & Hützen, H. in *Analytical Chemistry Symp. Ser.* Vol. 7 (ed. Frigerio, H.) 205–218 (Elsevier, Amsterdam, 1981).

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 500 words and the briefest of replies, to the effect that a point is taken, should be considered.

Taking on the Creationists

Philip Kitcher

Science and Creationism. Edited by Ashley Montagu.
Oxford University Press: 1984. Pp.415. Hbk £22; pbk \$9.95.

ON 5 January 1982, William Overton, justice in the Eastern Arkansas district of the United States District Court, delivered his decision in the case of *Maclean versus the Arkansas Board of Education*: "The relief prayed for is granted". So ended another phase in the long history of controversy about evolutionary biology and its place in science education. The attempt to introduce so-called "Creation science" into the American high-school biology curriculum by law foundered on the rock of Overton's decision. Much relieved, the prominent scientists who had led the resistance could return to work.

Two years and several volumes later, we have a new anthology, with nineteen articles, an editorial introduction and a historical preface. It is natural to wonder whether there is a niche for Ashley Montagu's collection to fill. True, Montagu reprints several pieces that are worth reading: Overton's decision, Roger Lewin's account of the notorious Arkansas Act 590 (which serves as the historical preface), and popular articles by Stephen Jay Gould and Isaac Asimov. It also contains some amusing celebrations of the great victory in Little Rock (by Michael Ruse and Gene Lyons, respectively). Yet, the question remains: is there a need for another volume on Creationism?

In my judgement, the answer is "Yes". Opposition to evolution and to the teaching of evolution continues in the United States, its traditional haven, and the Institute for Creation Research periodically claims that it is making headway elsewhere (notably in Britain and Scandinavia). Creationists seem to have returned to their favourite tactics — ambushes of local educators and assaults on textbook publishers. Until there is a body of literature that makes clear to those who make educational decisions, and to the public they represent, the magnificent accomplishment of Darwin and his successors, we shall need people to write about Creationism. Any anthology that promotes broad understanding of the overwhelming evidence for evolution and the blunders, distortions and pretences of "Creation science" should be welcome.

Unfortunately, Montagu's anthology fails to respond to the genuine need. It is a curious mixture. Relatively technical articles sit cheek-by-jowl with popular pieces, historical narratives are joined with sociological speculations and personal reminiscences. But it would be wrong to think of the book as a multi-faceted account, combining the insights of many

disciplines in a systematic way. There is nothing systematic about it. What we have is a loose collection of essays that do not gain through juxtaposition.

Still, the volume contains some good material. Besides the essays to which I have already referred, there is a characteristically informative introduction by Montagu

What was at stake in Little Rock?

The principal requirement of Act 590, passed by the Arkansas State Legislature in March 1981 and declared to be unconstitutional in January 1982 because it violated the separation of church and state, was that: "public schools within this state shall give balanced treatment to creation-science and to evolution-science. Balanced treatment of these two models shall be given in classroom lectures and in textbook materials taken as a whole for each course, and in library materials and other educational programs".

"Creation science" was defined as the scientific evidence for creation, and inferences from that evidence, indicating:

- Sudden creation of the Universe, energy and life from nothing.
- The insufficiency of mutation and natural selection in bringing about development of all living kinds from a single organism.
- Changes only within fixed limits of originally-created kinds of plants and animals.
- Separate ancestry for man and apes.
- Explanation of the Earth's geology by catastrophism, including the occurrence of a worldwide flood.
- A relatively recent inception of the Earth and living kinds.

and a brief, but intriguing, article by Gunther Stent, who argues that Creationism may do some good if it trims the excesses of "hyperevolutionism", represented most prominently by trends in human sociobiology. Two of the most diligent anti-Creationists contribute excellent essays. Laurie Godfrey's article (a revised version of a piece that originally appeared in *Natural History*) provides a lucid exposé of Creationist scavenging. Godfrey demonstrates how the words of scientists are ripped out of context and hammered into weapons that Creationists can use in debate. Even more valuable is Kenneth Miller's patient and comprehensive reply to Creationist arguments about radiometric dating and the character of the fossil

record. Those who pray for public understanding of the Creation-evolution controversy could hardly ask for more than that Miller's article will be widely read. Its remarkable clarity should enable anyone to appreciate the important points at issue — and should show the scientific community how to take on the Creationists and win.

However unsystematic, an anthology whose articles were all of such calibre would be worth reading. This is not the case with the present collection. Indeed, some of the articles in it are likely to provide grist for Creationist mills. Robert Root-Bernstein tries to resolve the Creationism controversy with some muddled philosophy of science, apparently fashioned from trying to synthesize incompatible points of view. Kenneth Boulding addresses the difficult issues surrounding the compatibility of evolution and theism, but the conclusions he is led to in his speculative explorations are hardly likely to satisfy those theists who take their science seriously. Beverly Halstead has a shorter way with such questions. Creationists will have a field day with Halstead's blunt professions of scientific atheism and atheistic science. The pity of it is that the methodological and theological questions are indeed central to the Creation-evolution controversy. In dealing with them in so amateurish a fashion, this anthology does them and the scientific community a disservice.

One other article deserves special mention, Sydney Fox's detailed and technical review of "origin of life" experiments. Those readers for whom Gould, Asimov, Miller, Godfrey and many of the others write, will find this article impenetrable. Nonetheless, by showing in detail what can be achieved in the synthesis of biologically significant molecules, Fox disposes of a source of familiar Creationist objections. Yet we ought to ask if this is the right kind of response. Evolutionary theory is surely stronger today because of ingenious experiments in Florida and careful excavations in East Africa. But the theory has been taught in schools around the world for a long time, and it has been right to teach it because of the strength of the long-standing evidence in its favour. It is surely wrong to think of Fox and Leakey as shoring up a tottering edifice.

Perhaps the public needs less information about current research and more straightforward exposition of the reasons why Biblical Creationism was rejected long before Darwin and why Darwin and his contemporaries accepted descent with modification. Some of the contributors to this volume begin the necessary work, but to carry it out effectively would require a much more disciplined book. □

Philip Kitcher is Professor of Philosophy at the University of Minnesota, Minneapolis, and author of Abusing Science: The Case Against Creationism (MIT Press, 1982).

History by numbers

C.W. Kilmister

Number Theory: An Approach through History from Hammurapi to Legendre.

By André Weil.

Birkhäuser: 1984. Pp.375. SwFr.64, \$25.

YEARS ago books on the history of mathematics, especially those addressed to the general public, were written by second-rate mathematicians who had turned themselves into something less than second-rate historians. More recently things have improved greatly and there are first-rate historians producing works of real scholarship.

André Weil's *Number Theory* is different again and even better. For here is a great mathematician, who has had, as it happens, a lifelong interest in the history of his subject, writing on the history of that branch of it for which he is most famous. He does so, moreover, in a most scholarly way and yet making the material available, not perhaps to the "educated reader" who is to get from this an initiation into number theory, as the author suggests, but at least to those whose mathematical knowledge is no more than is common to scientists, and who know no number theory. The material gathered together here was originally presented as lectures at the Princeton Institute for Advanced Study.

The book's title is accurate yet misleading. The first chapter surveys the first 32 centuries in as many pages, ending with the material available to Fermat (Euclid, Bachet's Diophantus and Viète's Diophantus). The extent of the mathematical learning of the author is clear:

Fermat. . . goes on to assign the work of Viète to 'geometry' and classifies that of Diophantus as 'close to geometry'. From our modern point of view, things look somewhat differently. Firstly, since so much of Diophantus, and even more of Viète, remains valid over arbitrary fields, we would classify this primarily as algebra. . . there is much in Diophantus and in Viète's *Zetetica*, which in our view pertains to algebraic geometry.

This sets the tone for the book. There is no truck with the purist tradition in the history of ideas that holds that the past must never be seen through the spectacles of the present. The history is told through its consequences for later number theory and for algebraic geometry.

The main part of the book comes in the next two chapters, one on Fermat and one on Euler. Fermat is seen as occupied initially with elementary problems of divis-

ibility and of the expression of numbers as sums of squares, but coming at the end to much deeper Diophantine problems, essentially to the unit of a real quadratic field. Weil says that "A substantial part of Euler's arithmetical work consisted in no more, and no less, than getting proofs for Fermat's statements . . .", but, in this context, shows in his next chapter with what depth and forward-looking intuition Euler considered the multiplicative group of integers modulo N , elliptic integrals, continued fractions, the zeta-function and prime divisors of quadratic forms. Here the history is full and careful, and "Fermat's legacy" is traced through Euler down to Mordell in 1922. Each chapter has appendices giving mathematical background. The account of Fermat has one on Euclidean quadratic fields, on curves of genus 1 in projective spaces, on Fermat's double equations as space quartics and on his method of descent and Mordell's theorem. Similarly, the Euler chapter is followed by notes on quadratic reciprocity,

a proof from 1912 that every integer is the sum of (at most four) squares and the addition theorem for elliptic curves.

There is a final short chapter on Lagrange and Legendre, followed by appendices on Hasse's principle for Ternary Quadratic Forms, on a proof of Legendre about binary quadratic forms and a proof of Lagrange about indefinite binary quadratic forms. The final word about this stimulating book should come from the second appendix to the chapter on Euler: Weil notes that the 1912 proof about sums of squares would have been easily understood by Euler, and

perhaps with a little more effort by Fermat, whose algebraic skills still fell somewhat short of the required level. That it was discovered so late may serve as an encouragement to those who seek elementary proofs for supposedly sophisticated results. □

C.W. Kilmister is Professor of Mathematics at King's College, University of London, and past President of the British Society for the History of Mathematics.

Ecologists' talk

Michael J. Crawley

A New Ecology: Novel Approaches to Interactive Systems.

Edited by Peter W. Price, C.N.

Slobodchikoff and William S. Gaud.

Wiley: 1984. Pp.515. £56.95, \$59.95.

A New Ecology brings together the thoughts of a group of voluble and enthusiastic American ecologists, plus some from John Lawton of the University of York, divided into sections covering resources and populations, life history strategies, ecology and social behaviour, and organization of communities, with a concluding "Synthesis". There are details of several admirable long-term studies (for example Whitham *et al.* on plants as genetic mosaics, Frankie and Morgan on oak galls, Lawton on bracken insects, Dayton and Tegner on kelp beds and sea urchins), while Wilbur describes some exemplary field experiments on tadpoles and their predators.

Most of the contributors express dissatisfaction with current ecological theory, and with competition theory in particular (although a dissenting voice is raised in Colwell's entertaining essay). But it strikes me that the authors have argued themselves into a corner by placing herbivorous insects at the centre of their view of ecosystems — the world is green, and ecosystems in fact consist of resource-limited plants, enemy-regulated insect herbivores, and food-limited natural enemies and decomposers. Despite its naivety, and a wealth of counter-examples, this generalization is as good as most in ecology. It certainly allows us to say that the last place one would look for competition as the dominant force in

the structure and dynamics of communities is amongst the insect herbivores. If the emphasis were on plant populations, or on vertebrate herbivores or decomposer species, then surely the authors' reservations about competition theory would be less vehement.

The book is much stronger when describing field ecology than when dealing with "the appliance of science". Clearly these are authors with a mission for the improvement of (other people's?) science, and a certain fervour pervades the opening chapter and several of the discussions. The insistence on scientific rigour is refreshing, but the impact is diluted by a maternally and rather self-righteous tone. Overall, the style of presentation betrays long hours of practice in the waving of arms; we find authors who write "illation" when they mean "inference", and who prefer "paradigms" to "hypotheses". Another irritating aspect of the book is that, like bananas and number 19 buses, the references come in bunches; in one staggering burst of erudition, Istock showers us with no less than 61 citations in a single sentence.

Many of the chapters have a curiously anti-theory tone to them; to read that solution of problems in the *New Ecology* will require "the attention of minds unfettered by preconceptions of resource limitation, equilibrium, optimization, competition, and the like" (Wiens, p.427), suggests that an understanding of these ideas is unnecessary (or even dangerous)! In fact, if some of the authors understood these notions more thoroughly, the book might have been a good deal better. For example, Strong muddies the water by dredging up a set of ideas from the 1950s which he christens "density vague ecology and liberal population regulation". He seems to have convinced himself that "theoreti-

New in paperback

Stephen R.L. Clark's *The Nature of the Beast: Are Animals Moral?*, which first appeared in 1982. Publisher is Oxford University Press, price is £2.95. For review see *Nature* 300, 136; 1982.

cians" (presumably the likes of Robert May) do not understand that, in the real world, factors other than population density affect the rates of birth and death. This extraordinary assertion is coupled with a number of misconceptions about the nature of population regulation and the detection of density dependence. Strong views it as a serious flaw that deterministic models do not describe the stochastic variability of real populations. Yet apart from adding noise, the alternative models he suggests contribute little to our understanding, while at the same time clouding the fundamental predictions which the simpler models are able to make.

One could easily come away from this book in the belief that the New Ecology is to be a science without theory, and with no explicit, simple models of dynamical behaviour. Even where the role of quantitative theory is acknowledged, the plea seems to be for increased "realism" at virtually any price. However, to replace linear density-dependent functions in deterministic models by strongly non-linear functions in stochastic models, is only sensible if these increases in complexity expose qualitatively different dynamic behaviour. There is no evidence of this here.

An important theme addressed by several of the contributors — Wiens and Price among them — is the central role of resources in determining community structure and function. There are formidable practical problems in this area; for example, while it is sometimes quite straightforward to determine the *use* of resources (for example to see where a butterfly lays her eggs), it is extremely difficult to measure resource *availability*, and virtually impossible to assess the *suitability* of the resources as the consumer itself perceives them (for example to understand why certain host plants do *not* have eggs laid on them). Problems of measuring "resource quality" in a currency appropriate to the consumer species are fundamental to the plant-herbivore studies which form the basis of so many of the papers in the book. Sad to say, no new approaches to resource assessment are described in these pages.

The title *A New Ecology* suggests a treasure trove of novel ecological insights. While it is far from that, the book may fulfil one of its stated aims by "providing a source of discussion topics", although it certainly will not guarantee that students who are beginning a research career in ecology can "readily assess how to make a major contribution to its development" (p.vii)! On balance, students are likely to benefit more from the good chapters than they will suffer from the poor. A New Ecology which coupled the excellence of the field experiments reported here, with the profound improvements in quantitative theory which are so signally lacking, would be a potent force indeed. □

Michael J. Crawley is a Lecturer in Ecology at Imperial College, University of London.

Star-watching from Leiden

Pieter B. Bosma

De Leidse Sterrewacht: Vier Eeuwen Wacht bij Dag en bij Nacht.

By Gijsbert van Herk, Herman Kleibrink and Willem Bijleveld.

Uitgeverij Waanders/De Kler, Grote Markt 9, Zwolle, The Netherlands: 1983. Pp.159. Dfl.32.50.

ONE of the oldest and most renowned scientific institutions in The Netherlands is the astronomical observatory of Leiden University: De Sterrewacht. As part of the celebrations on the occasion of its 350th anniversary in 1983, *De Leidse Sterrewacht*, a handsome, large-format history of the observatory was officially released.

The subtitle of the book, which in translation reads *Four Centuries of Alertness by Day and by Night*, needs an explanation. The anniversaries of the observatory are usually counted from 1633, because in that year modest observing facilities on the roof of the university building were put into service. However, lectures in astronomy had been given since the official appointment of Willibrord Snel in 1609. This adds 24 years to the observatory's history, but it does not extend its span to four centuries; 26 years are still missing. These years will be found in the future. As Professor H. van der Laan, chairman of the observatory, points out in his preface, an agreement on the cooperation of British and Dutch astronomers, of great importance for De Sterrewacht, does not expire until the year 2009.

From this it may be seen that the book should not be considered as a purely historical document. It is also not historically orientated in the sense that equal attention is paid to each period described. The average number of pages dedicated to each year grows roughly exponentially with the observatory's age, with about half of the book describing the past sixty years. This imbalance is justified if one remembers that an earlier record was written in 1933 by W. de Sitter, a former director of the observatory.

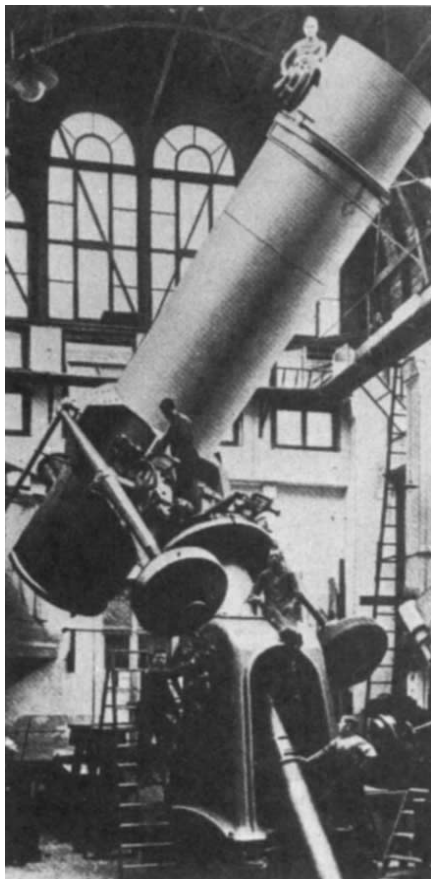
The first half of the book strongly reflects the efforts of successive directors to give a sound basis to the pursuit of astronomy in Leiden — in those days, too, it was sometimes difficult to find the necessary funds. In the second part the emphasis is on modern astronomical and astrophysical work. Although for many years now astronomy has been largely a matter of teamwork, the name of Professor J.H. Oort is prominent in the period after 1945. During this time the radio telescopes in Dwingeloo and Westerbork and the optical telescopes of ESO in Chile were built. Among other achievements, recent

years have witnessed a much better understanding of the structure of our Galaxy and completely new insights into the structure and behaviour of extragalactic systems. These are only two examples of successful scientific work; many more are briefly sketched in the book.

The authors G. van Herk and W. Bijleveld have written a vivid text in almost journalistic style, while H. Kleibrink has skilfully illustrated the book with a fine selection of many photographs and drawings. The text is in Dutch but a comprehensive summary in English together with a translation of the figure captions is given in the final 12 pages. Primarily, the book seems intended for those familiar with De Sterrewacht, but those interested generally in the development of astronomy may also like to read it. First, however, they'd probably better polish up their Dutch. □

Pieter B. Bosma is Senior Scientist in the Department of Physics and Astronomy at the Free University, Amsterdam.

150 years of astronomy



Big science at the turn of the century — assembly work at the Zeiss factory, Jena, on a 120-cm reflecting telescope for Babelsberg Observatory in Berlin. The illustration is reproduced from *The History of Astronomy from Herschel to Hertzprung*, by Dieter B. Herrmann, recently published by Cambridge University Press (price £12.50, \$24.95).

Improving muscular effort

R.C. Woledge

Handbook of Physiology: Section 10, Skeletal Muscle.

Edited by Lee D. Peachey, Richard H. Adrian and Stephen R. Geiger.
American Physiological Society
(distributed by Williams & Wilkins): 1983.
Pp. 688. \$145, £139.

THE American Physiological Society started to publish the series *Handbook of Physiology* 25 years ago. It was introduced with the following words:

The original literature of physiology has become so vast and is growing so rapidly that the retrieval, correlation and evaluation of knowledge has become with each passing year a more complex and pressing problem. Compounding the difficulties has been the inevitable trend towards fragmentation into smaller and smaller compartments, both of knowledge and of research skills. This trend . . . must be accompanied by the development of mechanisms for convenient and reliable reintegration in order that knowledge shall not be lost and research effort wasted.

Since then the problems have not diminished, and the publication in this series, for the first time, of a volume devoted to skeletal muscle has been eagerly awaited. To what extent can the editors and authors solve the immense problems of "reintegrating" this highly fragmented field, and present us with a "critical and comprehensive account of the physiological knowledge and concepts"? Inevitably the effort has to be made by a large team — in this case of 29 authors — and the resulting volume shows all the virtues and faults common to such works.

The advantage that stands out above all is the undoubted authority of each contributor within his field, and the corresponding confidence and enthusiasm with which they survey their chosen areas. Nearly all the chapters are strong, stimulating reviews. My particular favourites are those by Haselgrove (on X-ray diffraction studies), B. Eisenberg (quantitative ultrastructure), Gonzales-Serratos (inward spread of activation), Kushmerick (energetics) and by Peachey and Franzini-Armstrong (structure and function of membrane systems). All of these and many others succeed in providing a historical perspective and at the same time conveying something of the continuing excitement of developing areas; they are comprehensive and detailed but seldom boring.

The main fault of the book stems from the fact that the editors appear to have allowed their contributors undue freedom in choosing what they will cover. The consequence is that there are some substantial areas of overlap and some surprising — and serious — omissions.

For example, although 180 pages are

devoted to the structure of muscle, I could find no discussion of the lengths of the thick and thin protein filaments which make up the myofibrils. This question is crucial because the interpretation of the relation observed between muscle length and the force it produces requires a knowledge of these lengths. Worse still, the interpretation that is given uses filament lengths that have long been known to be somewhat wrong.

There is also no detailed discussion of the way in which actin activates the splitting of ATP, although the subject is touched on in two chapters, one of which deals in considerable and fascinating detail with the splitting of ATP by myosin alone. But a knowledge of what actin does is surely essential to bring together the biochemical study of myosin as an ATPase with the physiological study of energy transduction in muscle. A further criticism is that although the longest chapter in the book is devoted to the mechanism of calcium transport by the sarcoplasmic reticulum, no author has made it his business to

discuss the process of relaxation in muscle and the extent to which calcium transport and other processes contribute to and explain the dynamics of relaxation.

The weakest section of the whole book is undoubtedly the chapter on the mechanics and theories of contraction. It is much shorter than other contributions dealing with topics of corresponding importance and therefore lacks the detail and the authoritative tone which should be exhibited here; this chapter seems to belong in an altogether more modest textbook. It is particularly unfortunate that such a lacklustre performance is apparent at what should be the core of a well-balanced treatment of muscle physiology.

Yet, overall, there can be no doubt that the volume as a whole is a great success. It should become a standard work of reference which will indeed help to ensure that "knowledge shall not be lost and research effort wasted". □

R.C. Woledge is Reader in Experimental Physiology at University College London.

Meteoric principles

David W. Hughes

Physics of Meteoric Phenomena.

By V.A. Bronshten.
Reidel: 1983. Pp. 356. Dfl. 185, \$74, £46.95.

METEORS, commonly known as shooting or falling stars, are produced when small interplanetary dust particles hit the Earth's upper atmosphere and burn out in it. Incident velocities range from 11 to 74 km s⁻¹. The causative particles have masses between 10⁻⁶ and 10⁸ g. Beneath the lower limit the resulting ionization and excitation is under our detection threshold — these particles are retarded before massive ablation sets in and then drift down to the Earth's surface. The upper limit is somewhat arbitrary, being governed by the fact that on average only one 10⁸ g meteoroid hits the Earth each year. Also a 10⁸ g meteoroid encounters about its own mass of air during its atmospheric traverse. Larger bodies punch through the atmosphere like bullets and form craters in the Earth's surface.

The simplistic approach to meteor physics relates the observables, such as luminosity as a function of distance along the trail, beginning height and end height, deceleration, time of occurrence and spectra to the fundamental properties of the incoming particle — mass, velocity, density, composition and orbit. But nature is far from simple. Meteoroids are loose conglomerates — fragile dust balls which fragment in flight — and the relationship between their loss of kinetic energy and their physical parameters is complex. Large meteoroids, below heights of 80 km, are surrounded by shock waves and analysis of their motion has to be tackled by gas-

dynamical approaches. No longer are individual molecular impacts solely responsible for the melting and ablation of the meteoroid. Convective transfer, radiation and electronic heat conduction have to be taken into consideration.

Monographs on meteoritic phenomena are rare, the last two to be published being *The Physical Theory of Meteors and Meteoric Matter in the Solar System*, by B. Yu Levin (Izdatel'stv. Akad. Nauk. SSSR, 1956) and *Physics of Meteor Flight in the Atmosphere* by E.J. Opik (Interscience, 1958). Obviously the subject needs bringing up to date and V.A. Bronshten has done just that. He has assumed that the reader starts off with a good grasp of physics and applied mathematics but little knowledge of meteor science. This he builds up from first principles.

The Russian version of the book came out in June 1980 and the English translation was completed two years later, the hiatus being bridged by the addition of a short ten-page appendix. Bronshten covers the subject in commendable detail, and after developing the fundamental equations deals with the heating of the meteoroid, its ablation by vaporization, the spraying away of the molten layers and the formation of a shock wave. He then considers the luminosity and spectra of meteors, and how the curve of growth method leads to an understanding of chemical composition. Finally, there is treatment of meteor ionization, mass and density determinations, and the problems induced by fragmentation.

Bronshten's monograph is clearly written and well translated. It is a worthy addition to the few volumes on this rather specialized subject. □

David W. Hughes is a Lecturer in Physics and Astronomy at the University of Sheffield.